

Perceptual Control of Interceptive Timing

James R. S. Tresilian

Doctor of Philosophy
University of Edinburgh
1991



Declaration:

I declare that this thesis was written by me and that the research reported herein was conducted by me, unless otherwise indicated.

October 1991.

Abstract

The question of how actions involving the interception of moving objects are perceptually timed is addressed. This question has been intimately bound up with a debate which sets the "ecological" approach to perception and action (due to Gibson) in opposition to approaches which employ computational concepts. It is argued that modern versions of the two types of approach are not, in fact, opposed but are largely complementary and frequently equivalent. A general approach for tackling problems of perceptuo-motor control in humans and animals which integrates the two approaches is outlined. The problem of how interceptions of moving objects are perceptually controlled is investigated according to this general approach. First, the informational requirements of interceptive actions are analyzed. It is concluded that "time-to-contact" information is critical for accurate timing. The hypothesis, due to Lee, that animals and people assume the relative velocity between target and interception point to be constant when computing time-to-contact is discussed. A scheme for the continuous control of interceptive timing based on this strategy is formulated. Having established how time-to-contact information might be used to control interceptive timing the question of the perceptual source of this information is examined. A mathematical analysis of the visual stimulus is provided which clarifies and extends Lee's theory concerning the visual source of time-to-contact information. First, new sources of perceptual information about time-to-contact are described and second, a detailed analysis is presented of the conditions under which the optic variable tau, introduced by Lee, can play a role in the visual perception of time-to-contact. Several problems with using this variable are identified and ways of overcoming these problems are presented. The extensions to Lee's theory predict that certain perceptual variables in addition to tau play a role in controlling the timing of interceptions. Whether or not human subjects actually make use of these variables in timing interceptions is examined in a series of experiments. These experiments involved timing interactions with self-luminous objects in the dark. This allowed for control over exactly what visual information was available to subjects. The results suggest strongly that the variables predicted by the theory are actually used in interceptive timing tasks and that the perceptual source of some of these variables need not be vision. Finally, the implications of the theory and results for the perception of time-to-contact are discussed and several further research questions are identified and experimental techniques for investigating them outlined.

Acknowledgements

The research reported in this thesis was carried out at the Department of Psychology, University of Edinburgh during the years 1986-1989 and supported by a grant from the Medical Research Council which is gratefully acknowledged. I would like to thank Nick Chater for invaluable discussions which helped formulate the arguments presented in chapters 1 and 2: his intelligence and enthusiasm were a frequent source of inspiration. I am grateful to Professor Mike Morgan who allowed me a degree of freedom during the two years I worked for him without which it would surely have taken me far longer to finish writing up. Thanks are also due to Geoff Baldwin for designing and building the electrical circuits involved in measuring the speed of the balls in the experiments reported in chapter 7, Ewan Kirk for help with text formatting and Duncan Rand for help with data collection. Finally, I must express my gratitude to my supervisor, Dave Lee, for his help and insightful criticism throughout the course of this work.

Edinburgh, September 1991.

Table of Contents

Chapter 1

Introduction: Perception and action 1

- 1.1 INTRODUCTION AND OVERVIEW OF THE THESIS
- 1.2 DIRECT AND INDIRECT PERCEPTION
- 1.3 PERCEPTION AND ACTION
- 1.4 CONCLUSIONS

Chapter 2

Perceptual Control of Behaviour 27

- 2.1 INTRODUCTION
- 2.2 CONTROL
- 2.3 CLOSED LOOP CONTROL
- 2.4 PERCEPTUAL CONTROL: FIRST EXAMPLE
- 2.5 PERCEPTUAL CONTROL: SECOND EXAMPLE
- 2.6 PERCEPTUO-MOTOR CONTROL
- 2.7 CONCLUSIONS

Chapter 3

Control of Interceptive Timing 58

- 3.1 INTRODUCTION
- 3.2 INTERCEPTIVE ACTIONS
- 3.3 HITTING AND CATCHING
- 3.4 THE CONSTANT VELOCITY STRATEGY IN INTERCEPTIVE
TIMING
- 3.5 FORMULATION OF THE CONSTANT VELOCITY STRATEGY
- 3.6 CONCLUSIONS

Chapter 4

Visual Stimulation and Information 70

- 4.1 INTRODUCTION
- 4.2 AVAILABLE VISUAL STIMULATION
- 4.3 INFORMATION IN THE VELOCITY FIELD
- 4.4 TIMING INFORMATION DURING LOCOMOTION
- 4.5 INFORMATION ABOUT MOVING OBJECTS
- 4.6 CONCLUSIONS

Chapter 5
Obtaining Timing Information 98

- 5.1 INTRODUCTION
- 5.2 PROBLEMS IN EXTRACTING LOCAL TAU
- 5.3 PROBLEMS IN EXTRACTING GLOBAL TAU
- 5.4 MULTIMODAL AND INTERMODAL TIMING INFORMATION
- 5.5 CONCLUSIONS

Chapter 6
Empirical Investigations and the Distance+Velocity Hypothesis 118

- 6.1 INTRODUCTION
- 6.2 WHAT TIMING INFORMATION IS USED?
- 6.3 THE DISTANCE+VELOCITY ACCOUNT OF T_c PERCEPTION
- 6.4 EMPIRICAL STUDIES
- 6.5 IDENTIFYING EMPIRICAL PROBLEMS

Chapter 7
Experimental Investigations 133

- 7.1 INTRODUCTION
- 7.2 INVESTIGATION OF THE EFFECT OF VISIBLE SURROUNDINGS ON INTERCEPTIVE TIMING
 - Experiment 1
 - Experiment 2a
 - Experiment 2b
 - Experiment 3
- 7.3 INVESTIGATION OF THE EFFECT OF VISION OF THE INTERCEPTION POINT ON INTERCEPTIVE TIMING
 - Experiment 4
 - Experiment 5
- 7.4 CONCLUSIONS

Chapter 8
Conclusions 191

- 8.1 OVERALL SUMMARY
- 8.2 ADDITIONAL IMPLICATIONS AND FURTHER RESEARCH
- 8.3 CONCLUSIONS

| | |
|--------------------------------------|------------|
| Appendix A | |
| Accelerative Approaches | 203 |
| References | 206 |
| Publications | 219 |

1

Introduction: Perception and Action

§1.1. INTRODUCTION AND OVERVIEW OF THE THESIS

Activities involving interaction with objects or surfaces in motion relative to the performer, which we will later refer to as *interceptive actions*, frequently require precise *timing* if they are to be executed effectively. Examples of such actions include catching or hitting a moving object, placing the feet whilst running, plunge diving and landing from a fall. In this thesis the problem of how these actions are controlled is examined and the question of the informational requirements for the precise control of timing is addressed in detail.

Much of the recent research into the perceptual control of interceptive action conducted in psychology has been influenced in one way or another by the work of J. J. Gibson and those who have adopted his theoretical framework. This has been especially true of research concerning perceptual timing: analysis of the problems of perceptual timing by Lee and others has become the paradigm example of Gibson's strategy for the study of perception and action (Lee & Reddish, 1981; Turvey & Carello, 1986; Turvey, Shaw, Reed & Mace, 1981). Presumably because of this, perceptual timing has become a testing ground for Gibson's theory of "direct" perception: much of the published empirical research on the subject has concerned itself to a greater or lesser extent with this issue (e.g., Cavallo & Laurent, 1988; Cavallo, Laya & Laurent, 1986; Hofsten, 1987; Groeger & Brown, 1988; McLeod, McLaughlin & Nimmo-Smith 1985; McLeod & Ross, 1983; Savelsburgh, Whiting & Bootsma, 1989; Schiff & Detwiler, 1979; Schiff & Oldak, 1990). The results of this research have been inconclusive concerning the directness or otherwise of (visual) perception. Despite the obvious importance of this debate in the literature on perceptual timing, it has not significantly clarified the problems of how interceptive acts are controlled. In this thesis no attempt will be made to make an empirical contribution to the issue of direct perception. The reasons for not making such an attempt are detailed in this chapter. It is argued that although in the past it was meaningful to distinguish Gibson's approach from the standard approach to perception such a distinction can no longer be made. It is argued further that the "computational" approach to perception, usually considered as being due in

large part to David Marr, cannot be usefully distinguished from Gibson's approach thus making empirical distinctions impossible. The conclusions of the first chapter lead to the proposal that research into the perceptual control of timing should concentrate not on attempting to test the notion of direct perception but on solving the problem of how interceptive acts are timed. Chapters 2 through 6 then go on to offer a detailed analysis of the problems involved in timing interceptive action.

In chapter 2 a computational¹ style framework for posing problems of perceptual control is developed. It is shown how the notion of a unitary perception-action system which has its origin in Gibson's ideas can be precisely formulated. It is further shown how the "dynamical systems" approach to motor control fits in as an important part of a "computational" style approach. In chapters 3 through 6 a computational analysis of the problems of timing interceptive actions is presented following the framework detailed in chapters 1 and 2. Starting in chapter 3 with an analysis of the control problems of perceptual timing, the problems of how perception can provide the information required for this control are examined in chapters 4 and 5. In chapter 6 existing empirical research is examined in the light of earlier discussion and new empirical problems are identified, one of which is investigated in the series of experiments presented in chapter 7. In chapter 8 some of the remaining empirical problems are discussed in detail and methods for investigating them suggested.

§1.2. DIRECT AND INDIRECT PERCEPTION

At the time Gibson was developing his ecological approach to perception, received wisdom in psychology and in philosophy held that some sort of knowledge based inferential (cf Helmholtz's 'unconscious inference') or hypothesis testing procedures (Gregory, 1966, 1974) are required to generate perceptions from the insufficient data provided by the senses. So entrenched was this point of view that psychologists took the existence of such knowledge-based procedures to be self evident and the doctrine went virtually unquestioned. It is not surprising, therefore, that when Gibson proposed that no generative knowledge-based procedures are required in

¹ The term "computational" is here used to denote an approach which identifies different levels of understanding. The term was adopted by Marr for his approach to vision and is a generally appropriate (and convenient) term for a levels-based approach to understanding any functional process involving information processing as will become clear in the discussion presented in this chapter.

perception considerable debate ensued, a debate in which Gibson's approach is characterised as *direct perception* and the traditional approach as *indirect perception* (cf Micheals & Carello, 1981; Shaw & Bransford, 1977).

In this section it is argued that this debate can no longer be considered significant. It is shown that what Gibson offered was a simple rephrasing of one of the central problems of perception which allowed a conception of perceptual mechanism that does not require explicit inferential or hypothesis-testing procedures employing 'stored' knowledge. It is further argued that although Gibson's direct perception stands in stark contrast to the traditional conceptions of psychology it is not, in any sense that has been clearly stated, at odds with what has become the dominant paradigm of modern perceptual research, the 'computational' approach, as championed, for example, by the late David Marr (Marr, 1982). Five issues will be discussed: (1) the informational sufficiency of sensory stimulation; (2) whether perceptual illusions can be considered counter-examples to the informational sufficiency of the stimulus input; (3) what the role of information processing is within an approach that takes the stimulus to be informationally sufficient; (4) whether the mechanisms of perception can be considered as computational; and (5) levels of explanation in a computational understanding of perception.

(1) Informational Sufficiency

The traditional conception of knowledge-mediated perception is founded on the observation that the perceptual input (the stimulus) contains insufficient information for veridical perception of the environment. In the case of vision this is particularly clear: any given (retinal) image — a 2-D pattern — could have been generated by an indefinitely large (possibly infinite) number of different 3-D physical configurations of reflecting surfaces. This remains true if a time varying image is considered — there are still fewer dimensions in the stimulus than in the environment. The logic here is impeccable — the stimulus does not logically determine the environment (projections from a space of $n + 1$ dimensions to one of n dimensions do not admit unique inversion). Thus, we have an argument which runs as follows:

(1) Stimulus patterns are typically consistent with a large (infinite) number of logically possible environmental states of affairs (objects, events, etc.) and hence the stimulus does not uniquely inform about the environment — the stimulus is informationally impoverished.

(2) Because of the situation described in (1), in order for veridical perception of the environment to be possible information must be added to the information provided by the stimulus². This information comes from sources internal to the perceiver and is based on past experience with the world (memory).

This is the essence of the argument upon which the traditional approach is based. It leads to the conception of perceptual mechanism as an active process of *adding* information from some internal store (memory) to the insufficient information in the stimulus. Such mechanisms explicitly use knowledge about the structure of the world/environment to reason out what the immediate environment must be given the evidence of the senses.

Interestingly, we can rephrase the above to give exactly the opposite characterisation of the informational sufficiency of the stimulus, thus allowing a completely different conception of mechanism. The traditional psychological approach asserts that veridical perception of the environment is possible if the information provided by the stimulus is supplemented by information about the structure of the environment — ‘sense data’ and ‘memory’ together provide sufficient information. Thus, if the structure of the environment is known (in sufficient detail) the stimulus input can be rendered unambiguous. This is the same as saying that since the environment is known to have a particular structure, then a given stimulus input must have been generated by a unique environmental configuration. That is, knowledge of the general features of an environment’s structure allows all except one of the possible interpretations of a particular stimulus input to be ruled out.

Suppose that the perceiver lives in a world which has certain structural constraints. Suppose further that given these constraints a particular stimulus pattern, *logically* consistent with an arbitrarily large number of different physical configurations, could only have been generated by one and only one configuration in the perceiver’s world. Under such conditions, it is not necessary to use knowledge of the world to rule out all the possible (but not actual) environments that could logically have given rise to the stimulus pattern. The logical state of affairs is quite irrelevant if the logically possible environments are impossible in the world inhabited by the perceiver. The perceiver hardly needs to rule something out if it could never exist in the first place. Thus, if the structure of a perceiver’s world is sufficient to render

² Note that information from one stimulus modality added to information from another modality is not what is being considered here since this may obviously be considered part of the stimulus input (consider Gibson’s discussion of ‘intermodal invariants’).

the stimulus input unambiguous (the traditional account supposes this), then for the perceiver living in this world the stimulus itself is unambiguous. To put it another way: given that a perceiver (organism) lives in the environment that it does, the fact that a given stimulus pattern could *logically* have been generated by all sorts of physical configurations is irrelevant if, in the perceiver's environment, it could only have been generated by one configuration. Looking at the situation in this way, it may be argued that given the constraints of the real world perceptual stimuli are, for most of the time, informationally sufficient.

This alternative way of looking at the situation was identified by Gibson and clearly leads to the rejection of the notion of perceptual mechanism outlined earlier — if, in the real world (as opposed to logically possible worlds), the stimulus is unambiguous, then procedures to disambiguate it are redundant. Gibson's rejection of knowledge-mediated perceptual mechanism led him to suggest that what perceptual mechanism was doing if not adding information was detecting or 'picking-up' information (Gibson, 1966). Neither Gibson nor his followers ever provided a model of a perceptual information picking-up mechanism. However, Gibson did suppose, rather vaguely, that perceptual processes were closer to the pick-up of radio stations by radio receivers (Gibson, 1966; Micheals & Carello, 1981) than to processes of inference and hypothesis testing. It was this kind of analogy that led to the idea that perception involved a process of 'resonance' to stimulus information (Gibson, 1966) and perhaps also to the out and out rejection of any kind of information processing or computation in perception by Gibson and others (Gibson, 1966, 1979; Shaw & Bransford, 1977; Turvey, 1986).

It is this rather obscure attitude towards mechanism that is the source of most real controversy that remains concerning Gibson's ideas since the notion that the stimulus can be informationally sufficient under environmental constraint is now uncontroversial as amply demonstrated by recent work on low-level processes in vision, for example, stereopsis (Marr & Poggio, 1976, 1979; Mayhew & Frisby, 1982; Mayhew, 1983); colour (Land & McCann, 1971; Land, 1986; Marr, 1982); both discrete and continuous visual motion (Aloimonos, 1988; Longuet-Higgins & Prazdny, 1980; Koenderink, 1985, 1986; Ullman, 1979, 1983; see also chapter 3) and binocular visual motion or rate of change of disparity (Aloimonos, 1988; Regan & Beverley, 1979; Waxman & Wohn, 1988). A significant portion of this work is devoted to finding structure in the stimulus which informs uniquely about en-

vironmental structure given identifiable environmental constraints (Marr, 1982). For example, in a recent review of the computational approach to the perception of three-dimensional environmental structure from image motion Ullman observed that, “[t]he main conclusion has been that when the changing image is induced by rigid objects in motion, the 3-D structure of the objects is determined uniquely by the 2-D transformations of the image.” (Ullman, 1986; pp. 17). In fact, from the computational point of view the traditional position can look bizarre; for example Prazdny (1980) wrote,

“It is surprising that Gibson’s ... views have been attacked ... so often in view of the existence of other ... far more radical and eccentric theoretical standpoints (Gregory, 1972, 1979; Oatley, 1978). Surely these views ... are much more radically wrong than Gibson’s.” (pp. 394)

The existence of visual illusions has long been taken as providing difficulties for Gibson’s view that the stimulus is informationally sufficient. If what we have said about the computational theory is correct then illusions pose the same difficulties for this approach as for Gibson’s. The problem of illusions is discussed below.

(2) Perceptual Illusions

The existence of illusions has often been taken to show that Gibson’s approach is unworkable in general. Illusions are variously taken to contradict the Gibsonian approach in that they suggest that the environment is not uniquely specified by the perceptual input or that “constructive” perceptual processes are in operation (e.g., Gregory, 1974, 1980). Gibson is able, however, to give the following account of illusion.

As noted earlier, if a perceiver’s world has sufficient structure, then it is possible for the stimulus to inform uniquely about environmental properties, events or states of affairs. In this case, if the stimulus has a particular structure, *S*, then there will be a unique environmental situation, *E*, that could have given rise to *S*. *S* may then be said to *specify* *E* — *S* can be used as *information* about *E*. Suppose that an organism exploits the fact that *S* specifies *E* in the world it inhabits. If this organism perceives *E* when *S* is detected then supplying it with *S* will lead to it perceiving *E*. Suppose now that the organism is placed in an artificial environment

with different structural constraints such that S is no longer associated uniquely (or even at all) with E. In such circumstances if S is presented E will still be perceived even though E may not now be present. The perception of E in this case is illusory. This kind of account has been presented consistently by proponents of direct perception (e.g., Turvey et al., 1981). Gibson argued that such illusory perception may occur in the laboratory where the structural constraints which hold in the natural world may be deliberately violated (impossible figures), the stimulus input may be impoverished (target stimuli are viewed in restricted contexts such as in the dark), or the subject's normal behaviour during perception is prevented (the head is held still, for example). In these cases the available information may specify something which does not correspond precisely with physical reality. This line of argument appears to be appropriate for illusory perceptions such as seeing a straight stick as bent in a glass of water or a 3-D object when presented with a hologram. Illusions of this kind are informational — they would be perceived by any perceptual system which exploited the relevant dependency of stimulus structure upon environmental structure like that between S and E above.

The above argument deals with the objection that illusory perceptions are inconsistent with the notion of informational sufficiency. In fact, it is possible to test whether a perceptual system is relying on the existence of a particular physical constraint by presenting a subject with perceptual displays which violate this constraint. If the constraint is being exploited then illusory perception of displays which violate the constraint are predicted — the perceptual system will treat the display as if the constraint held. Research in computational vision routinely uses this technique (see, for example, Marr, 1982; Ullman, 1979)

There are certain illusions which may not be adequately accounted for by the above considerations. These are illusions in which there is apparently no information missing (no impoverishment) and no violation of physical constraints. Such illusions are exemplified by many of the "standard" visual illusions. For example, "twisted cord" illusions where people perceive illusory spirals rather than the concentric circles actually in the figure (Fraser, 1908); illusions where straight lines are seen as bent or curved such as the directional illusions of Herring and Wundt (see e.g., Luckiesh, 1965); illusions where parallel lines are seen tilted towards or away from one another as in the Zöllner illusion or the Münsterberg or "café wall" illusion; or illusions of size and length in 2-D pictures such as the Müller-Lyer

illusion. Illusions of this kind have often been taken as evidence for constructive or hypothesis-testing processes in perception (e.g., Gregory, 1974).

Most illusions of this kind need not be evidence of interpretative knowledge based procedures, however. It seems reasonable to suggest that these illusions fall into two classes depending on the kind of explanation that best accounts for them. As an example of the first class, consider illusions in which pictures of objects which are objectively the same size are presented in a context such that they are perceived as differing in size. Perspective drawings often have this characteristic. In such cases the visual system has potential access to the fact that the pertinent items are of equal size (their retinal images are the same size), it fails, however, to detect this fact. In perspective drawings where two figures of the same size are seen as different, it may be supposed that there is information present which specifies that the two objects represented are at different depths. Equal image size in this context specifies that the objects are of different sizes (the "nearer" is the smaller). The visual system is required to tell a person about the sizes of objects and is configured to detect the relevant information. It is not required (except perhaps in perceptual experiments) to inform about the size of retinal images. Therefore, in the case we have just described, the perceiver has potential access to information about the veridical situation but is not able to use it. It could be said that the visual system is set up to answer certain (functional) questions about the world. It cannot be expected to supply a sensible answer to any question that might be posed of it.

The second type of explanation proposes that the perceptual mechanism is responsible for the illusion but not because it is testing hypotheses or adding information. In this case the illusion is not due to the fact that the perceptual system is detecting one kind of information and failing to detect another kind as in the last example. Rather it is the method by which the information is being extracted that is responsible. Mechanism based accounts of visual phenomena have a long history — an explanation of Mach bands based on lateral inhibition was provided by Mach himself. Morgan has recently provided an account of this kind for the twisted cord and Münsterberg illusions (Morgan & Moulden, 1986; Morgan & Hotopf, 1989. See also Grossberg & Mingolla, 1985). The idea behind Morgan's account is that information about certain large scale environmental features, which have relatively large retinal images, is extracted through what amounts to the integration of small

local measurements on the image. For example, if tilts are detected by local measurements along a “border” in the image, the perceptual effect will be of a global tilt. This kind of local process will typically extract global tilt quite effectively and be robust against partial occlusion as Morgan and Hotopf remark. However, it is possible to devise visual displays in which there are local tilts along a line but globally there is no tilt at all. Under these conditions the visual system detects a global tilt in the line; Morgan and Moulden have suggested that this is the case for standard twisted cord illusions and (derivatively) the Münsterberg illusion.

Direct perception allows accounts of illusions of the first two kinds described in this section but does not appear to allow accounts of the third kind. Discussion of information extraction is explicitly eschewed in direct perception and hence explanations of this kind are excluded *a priori*. Gibson admits that perception has a mechanism so if mechanism based accounts are excluded such exclusion is somewhat arbitrary. There does not seem to be any compelling reason for denying the possibility that illusions exist which find their explanation in the workings of the perceptual mechanism. Clearly, the computational approach can deal effectively with the possibility of such illusions.

It is clear, therefore, that the notion of informational sufficiency is quite compatible with visual illusions; indeed, under certain circumstances illusions are predicted. Given this and that it is generally accepted that the stimulus can be informationally sufficient and that information need not be added from memory, what kind of perceptual processing is required? Gibson, as mentioned above, envisaged some kind of information pick-up process with perceptual mechanisms “resonating”, in some undefined way, to stimulus information. The computational approach is much more precise about what kind of perceptual processing is required as explained below.

(3) The Role of Perceptual Processing

The role of perceptual processing as it appears in the computational approach is to transform information present in the stimulus input(s) in a raw and *unusable* form into a *usable* form (it may be used to control and coordinate actions or as a basis for reasoning and planning). Marr (1982) adopted the term *implicit* for information in an unusable form and *explicit* for information in a usable form. Thus a perceptual processing module accepts as input an ‘array’ of data in which certain required

information is only implicit and delivers as output an array of data in which this information is explicit³. Such a transformation may be considered to be computation or information processing. One typically calls an array of data (carrying certain information explicitly and other information implicitly) a *representation*; under such a definition the stimulus input itself (e.g., the retinal image) is a representation (Marr, 1982). Indeed anything may be viewed as a representation, Gibson's optic array and optic flow are also representations in this sense. To call something a representation is simply to regard it from a functional perspective — to treat it as the input to (or output of) a computational process.

As an example of implicitness versus explicitness of information consider stereopsis. The images projected onto the receptive surfaces of two eyes or cameras carry, to a scale factor (the interocular separation), information about the depths of points on the visible surfaces which give rise to structure in the images. In this form (carried in two separate images) the depth information is unusable; in order to use the information the disparities in the images of the same physical points or surface features must be measured (computed) — the disparity is, to the interocular scale factor, explicit information about depth.

Much of the perceptual processing proposed in the models developed in the computational approach involves making information explicit by separating this information out from a representation in which it is mixed up with other information (see Marr, 1982). Examples of this appear in chapters 4 and 5 below. Interestingly, therefore, information processing in the style of the computational approach stands in stark contrast to that traditionally envisaged in psychology since to make information explicit a process of 'discarding information'⁴ is often required. The traditional view, it will be remembered, is of a process of adding information. In this way the processing of the computational mechanisms found in recent models of perception is not too far removed from the process of picking-up radio stations where the signal of a desired station needs to be separated out from other signals;

³ Following this idea modern theory concerning the function of visual cortex holds that it is involved in creating representations of differing degrees of abstraction which represent different information explicitly (Ballard, 1986; van Essen & Maunsell, 1983). Ballard, for instance writes that, "[d]ifferent cortical areas ... represent information at different levels of abstraction" (1986; pp. 73).

⁴ I place this phrase in quotes because information may not literally be discarded but simply separated out from other information; the perceiver may not lose it completely unless, of course, it is of no use to him.

some theorists even find the term 'resonance' appropriate (Grossberg, 1980; 1983).

As noted above, Gibson was extremely vague about what constituted the process of what he termed information pick-up and subsequent work by his followers has not appreciably altered this situation. The ideas described here seem consistent with Gibson's thinking and do succeed in providing an explicit description of what perceptual processing actually does which is specific enough to allow the development of formal and implementable models. The computational approach proposes algorithms for extracting information about properties of the environment. These algorithms only work in a world which satisfies the relevant constraints and they do not involve the addition of stored information from memory. The researcher requires knowledge of the constraints of the real world to design such an algorithm of course, but the algorithm itself does not contain this knowledge. It simply works because the constraints hold in the world in which it is designed to operate — try to use it in another "world" and it will probably fail. Current computational approaches arose in direct opposition to the knowledge driven approaches of early research in machine vision and artificial intelligence (Marr, 1982). In these earlier approaches knowledge of a particular visual domain such as what is typically to be found in an office is used to segment the image and recognise objects, such as telephones (Tenenbaum & Barrow, 1976).

Rather a lot has been said above about information without saying very much about what is meant by this term and it is probably worth making a few remarks regarding the usage of the term information at this point. The concept of information has been the subject of considerable debate and discussion and it has proved more or less impossible to provide a definition that meets with everyone's approval (for recent discussion see e.g., Barwise & Perry, 1983; Chater, 1989; Dretske, 1981; Turvey & Carello, 1985). Shannon's "information theory" (Shannon & Weaver, 1949) was founded on the conception of "information" that arose in nineteenth century physics, i.e., information as $H = -\sum p_i \log_2 p_i$ entropy. According to this view information is closely related to pattern or organization — it is in the order of a sequence of symbolic tokens drawn from a finite inventory or in the organisation of the states of a system that information is carried.

It is clear that pattern is more abstract than the physical medium that is itself patterned — the same pattern could be realised in all sorts of different physical

media — a triangle is a triangle regardless of what it is made of. It is this fact that underlies one of the most important aspects of information technology: the same pattern can be carried in sound, light, electrical “signals” and so forth, and translated between them by means of suitable transducers. For example, a microphone transduces acoustical pattern into electrical pattern and a loudspeaker transduces it back to acoustical pattern again. Since information is identified with pattern in information theory it is clear that information is abstract and independent of its physical realisation. In other words, we might say that information is intrinsically *amodal*. If the senses are considered as instruments for the pick-up or extraction of information as they were by Gibson, then we should not be surprised to find that the same information can be extracted by different senses (perceptual systems) as documented by research on sensory substitution devices (White et al, 1971).

This much is hardly controversial. The problem is that describing information as pattern (although useful quantitatively and technologically) says nothing concerning what the information is information about, i.e., what it *means*. Indeed, the usual understanding of the term information is ‘meaning’. Pattern itself does not mean anything but serves rather as the carrier of meaning, after all, the same pattern could be used to mean completely different things and *vice versa* — the same meaning could be carried by completely different patterns (think of different languages). Thus a distinction should be drawn between the quantitative information content of a pattern (its “bit” content) and the meaning of a pattern, its *semantic* content. In perception the structure of the environment causes perceptual media, such as the optic or acoustic arrays, to be structured and as Gibson pointed out, structure in perceptual media could be specific to environmental structure. In this way structure in, say, the optic array carries information about environmental structure — the presence of a particular optical structure implies the existence of a particular environmental structure. This idea that information about something is the same as specificity to something (where that specificity arises through natural causal processes) is perhaps the most useful conceptualisation of perceptual information that exists and is the one that will be adhered to here.

It will now be argued that there can be no argument concerning whether or not perceptual mechanisms may be regarded as computers processing information. It is proposed that to treat a system as a computer is simply to hold a certain kind of attitude towards that system.

(4) The Computational Attitude

Consider, for example, the relation of dynamical analogy between two physical systems (Olson, 1958). Two physical systems are said to be dynamically analogous under some description if they can be modelled using dynamical equations of the same form. For example, consider the following system of ordinary differential equations:

$$\begin{aligned}\frac{dx_1(t)}{dt} &= x_2(t) \\ \frac{dx_2(t)}{dt} &= ax_2(t) + bx_1(t) + cF(t)\end{aligned}\tag{1.1}$$

These equations can be used to model, for example, a mechanical system or an electrical system. Consider a linear damped mass-spring system, to which an external force $f(t)$ is applied. Suppose the mass on the spring is m , the stiffness of the spring k and the damping coefficient β . We may define $x_1(t)$ in (1.1) to be the displacement of the mass from its equilibrium position at time t , and $x_2(t)$ to be its velocity. Then (1.1) is a model of the mass-spring system if $a = -\beta/m$, $b = -k/m$, $c = 1/m$ and $F(t) = f(t)$.

Alternatively, we may consider the electrical circuit in which a capacitor of capacitance C , an inductor of inductance L , and a resistor of resistance R , are connected in series to a battery of electromotive force $E(t)$. In this case, we can define $x_1(t)$ in (1.1) to be the integral of the current flowing through the circuit, and $x_2(t)$ to be the current itself. Then (1.1) is a model of the electrical circuit if $a = -1/LC$, $b = -R/L$, $c = 1/L$ and $F(t) = E(t)$.

These two systems are said to be dynamically analogous, since they share a dynamical model of the form (1.1). By making the appropriate identification between the variables of the two systems, it is possible to learn about the behaviour of one by observing the behaviour of the other. This is no mere in principle possibility, but rather is the foundation of analogue computing (Wilkins, 1970). An (electronic) analog computer consists of a variety of electrical components which can be selected and interconnected (programmed) such that the resulting circuit is dynamically analogous to the system under study. Voltages measured at certain points within this electrical circuit can be used to represent the variables of the modelled system.

Any physical system which is dynamically analogous to some second physical system may be viewed as an analogue computer which simulates the behaviour of

this second system. One could use the mass-spring system in the above example as a computer to simulate the behaviour of the electrical circuit, or *vice versa* — either may be viewed as a computer, but neither need be. To treat the system as a computer is to hold a particular attitude towards that system—the system is not intrinsically a computer. The crucial feature of this attitude is the theorist's reaction to a departure of the observed behaviour of the first system and its mathematical (dynamical) description. If the system is treated purely as a physical system which is to be modeled mathematically, then if the mathematical model fails to describe the system's behaviour it is taken to be inappropriate and revised. On the other hand, when the system is treated as an analogue computer if its behaviour departs from that predicted by the mathematics, then the circuit fails to realise the appropriate mathematical object, and the circuit is duly repaired or reprogrammed. In so far as a physical system can be idealised by a mathematical object, the mathematical object may be viewed as describing the behaviour of the physical system (descriptive stance), or the mathematical object may be seen as being computed by the behaviour of the physical system (prescriptive stance). In short, to treat a system as a computer is to treat its mathematical idealisation as prescribing how it should behave rather than describing how it does behave⁵. Since it is performing an informational function, it is very weak claim to say that the visual system may be seen as a computer.

Thus, according to a sufficiently general notion of computation, perceptual systems may be treated as computers, under an appropriate idealisation. To say that a system is a computer places no constraints on its mechanistic/physical structure. In particular, it does not entail that the system effects any explicit symbol manipulation in the style of a von Neumann machine or that it explicitly make inferences or test hypotheses.

The style of computation used in most models of early perceptual processing in biological systems is of the 'parallel distributed' kind involving connectionist networks and is close to the analogue style (e.g., Arbib & Hanson, 1987; Ballard,

⁵ The idea that to talk of something as a computer is to adopt a prescriptive stance towards its mathematical idealization can be generalized to other devices. Anything that admits of a functional description (i.e., can be described as doing something for a 'purpose') is potentially able to go wrong—its function *prescribes* its behaviour and if the behaviour fails to generate that function the device is said not to be working (properly or at all). Any object that can be functionally described is a machine — in the case of certain kinds of informational function it is a computational machine or computer.

Hinton & Sejnowski, 1983; Grossberg, 1976, 1983; Hinton, 1980; Marr & Poggio, 1976; Poggio & Reichardt, 1976). It is very far removed from serial, logical processes working with explicit knowledge bases (memory). Indeed, Poggio & Koch (1985) describe some interesting analogue methods for solving certain computational problems in early vision. For example, they describe two chemical networks (sets of interrelated chemical reactions) which compute the smoothest velocity field. Chemical reagents input to the reaction system represent the measured velocities at points in the image and likewise, concentrations of chemical reagents in the reaction system represent the computed smoothest velocity field. Horn (1974) was one of the first to explicitly propose an analogue method for carrying out an early vision computation. He devised an electrical network (of resistors and current sources) which computed lightness from an image. These analogue networks may be viewed either as physical systems which may be modeled by a particular dynamical system (a set of equations such as 1.1 above) — the descriptive stance, or as computational devices constructed to perform a particular function — the prescriptive stance. It cannot be sufficiently stressed that the notion of computation is quite general — it is simply a term which becomes appropriate for system description when a functional attitude is adopted.

(5) Levels of Explanation

Marr argues that understanding a complex information processing system, such as a perceptual system, requires understanding that system at different “levels” (Marr, 1982; Marr & Poggio, 1977). Marr suggested that there are at least three levels at which it is necessary to understand a perceptual system: (1) the level of computational theory; (2) the level of algorithm; and (3) the level of hardware implementation. As Marr (1982) remarked, the computational level is close to the type of understanding sought in Gibson’s ecological approach. Gibson argued that one needs to understand what exactly is being perceived (what environmental quantities, properties, animal-environment relationships and so forth) and what perceptual information is available to specify these perceivables (and also, perhaps, what information an animal is actually using). This is what understanding at the computational level is all about. One seeks to understand what is being perceived and how the structure of the environment, in the form of physical constraints, allows the stimulus to carry information about these perceivables. Once sufficient constraints have been identified the stimulus specifies the perceived properties.

As discussed above, the ecological approach goes on to suppose that the information is “picked-up” by the perceiver. There is no further elaboration of this pick-up process, however, beyond claims that it does not involve computation and information processing. Indeed, Gibson and his followers explicitly place elaboration of the pick-up process outside of psychology (Gibson, 1979; Reed, 1980) and Gibson has been interpreted (Ullman, 1980) as denying the possibility of any intermediate level of explanation between the ecological level (\equiv computational level) and the level of the physiological hardware (or “wetware”). Under this interpretation it is the denial that there can be an intermediate level of explanation that sets the ecological approach apart from the computational approach. (The latter introduces the algorithmic level between the computational and implementational levels.) In line with the computational approach, it will be argued here that the algorithmic “level” is necessary if we are to understand a perceptual system fully. On the other hand, and in line with the ecological approach, it will be argued that the algorithmic level is not, in fact, a “level” at all in the sense that the computational and implementational levels are.

An algorithm is a method (a specified procedure) for obtaining from a set of low-level operations (basic operations) some more complex or higher-order operation. For example, a digital computer has a set of simple basic operations which the hardware of the machine is capable of performing (for an especially clear discussion see Harel, 1987). These operations simply involve turning microscopic “switches” on and off. The task of a computer programmer is to get the machine to perform some complicated operation like checking the spelling of words in a body of text or solving a set of mathematical equations. To do this he needs to be able to express the required operation of the machine in terms of the basic operations which the machine can carry out⁶. Once this has been achieved the machine can be made to perform the required overall operation simply by making it carry out the structured set of basic operations which express that operation. The expression of the high level operation in terms of the basic operations thus constitutes a means for obtaining the former from the latter. It is a method or algorithm for obtaining the high-level operation.

It is clear from the preceding discussion that an algorithm simply tells us how to

⁶ In the early days of computing the programmer had to do all the work involved in formulating such an expression. Today the task is made considerably easier due to the availability of high-level languages like PASCAL, FORTRAN or C.

obtain a high-level operation from some specified set of simpler (basic) operations. As such it is the analogue, in a functional context, of the following situation from physics. Suppose we have a physical system whose behaviour we can observe at a macroscopic scale and that it can be given a description in terms of macroscopic observables. A gas is a simple example: the behaviour of gases can be described in terms of macroscopic observables (such as pressure, temperature and volume) and various laws which govern these variables. The same gas can be described microscopically in terms of its constituent particles and the associated laws which govern the behaviour of these particles. In the latter description one finds no hint of pressure, temperature or volume and the associated macroscopic laws, yet both it and the macroscopic description describe the same gas. Clearly both descriptions are related and a major problem of nineteenth century physics was to discover how — how could one go from the microscopic description to the macroscopic description. The solution to this problem was provided (partially) by statistical mechanics, a theory which showed how it was possible to obtain a macroscopic description of a gas from its microscopic description. Here we have microscopic elements which do certain things (move through space, collide and exchange momentum) and a theory which shows how we can go from a collection of such elements to a system with pressure, temperature, volume and so on. Statistical mechanics shows how the two descriptions of a gas are related: without it we would not have a proper understanding of gases.

The situation is much the same in computation. We have a description of a system in terms of simple (microscopic) operations and a description of what we want it to do in terms of high-level (macroscopic) operations. A means is required to get from one to the other. Similarly, if we want to understand a machine carrying out a computation it is not enough to describe its overall operation and the operation of its microscopic components. There must also be a means for relating these descriptions. In computation such a means is called an algorithm. It should now be clear that an algorithm is not itself a level of description but simply a means for relating levels of description just as statistical mechanics is not itself a level of description of a gas. Thus a complete understanding of a computational system (of whatever kind) requires an understanding of algorithm. It should be noted, however, that the algorithms of connectionist networks and analogue machines need not be at all like the serial, step by step algorithms of digital devices.

Conclusions In this section it was argued that the differences between Gibson's direct perception and the mediated view of perception traditional in psychology and philosophy do not succeed in distinguishing Gibson's position from that which characterizes the computational approach. The computational approach is founded on the hypothesis that there is sufficient constraint in the world to render the perceptual input informationally sufficient and is thus consistent with Gibson's position. Instead of characterising perceptual mechanism as a process which adds information to an impoverished retinal image, the computational approach can be interpreted as conceiving of perceptual mechanism as a process which separates out (or even discards) information and acts to put information into a useable (explicit) form. It is hard to see that there can be any disagreement with Gibson in this matter. Finally, to view the perceptual mechanism as computational is simply another way of conceiving it as being functional. Gibson certainly would not have denied the functionality of perception and thus there cannot be any disagreement as regards this issue (not at any fundamental level at least).

One important aspect of Gibson's thinking which has not so far been discussed was his insistence that perception could only be considered in relation to action, something hardly ever touched on in the traditional psychological approach. Gibson's view was that perception delivered information about what could be done with the environment — its *affordances* (see e.g., Gibson, 1977). This point of view has also aroused controversy (e.g., Fodor, 1980; Fodor & Pylyshyn, 1981) even amongst those who have otherwise supported Gibson's position (Cutting, 1986). In the next section the relation between perceiving and acting is examined.

§1.3. PERCEPTION AND ACTION

Gibson repeatedly stressed the intimate relation between perception and action insisting that perception could only be understood when considered in relation to the functions it serves in telling an animal what it can do in the world. This point of view was not characteristic of the thinking of Gibson's contemporaries and has been often presented as radical and controversial (e.g., Reed, 1989; Turvey, Shaw, Reed & Mace, 1981; Shaw & Turvey, 1981; Turvey & Carello, 1986; Turvey & Kugler, 1984). Contrary to these presentations it will be argued here that Gibson's view is uncontroversial. Before embarking on this discussion it will be useful to be precise about exactly what "action" will mean in the following arguments since a

convincing case has been made indicating that many researchers take actions to be identical to particular movements of the body (see Kugler & Turvey, 1987; Newell, 1978; Reed, 1984, 1988).

Consider, as an example, the action of closing a door. It is obvious that this action is performed when a door, initially open, ends up closed. All we are interested when we refer to this action is the transformation from the door being open to it being closed — the interest is in *what* is done. This can be taken as the defining feature of action: when the term action is used one is interested only in *what* is done (the state or situation that is brought about) not in *how* it is done (the mechanism of the causal process that brings about that particular state or situation). In other words, when one speaks of an ‘action’ one is referring only to the *bringing about* of a particular state of affairs or the *effecting of a transformation* from some arbitrary initial state to a particular (final) state. If the particular final state is called the ‘goal’ then ‘action’ is synonymous with ‘achieving a goal’. This is, of course, the commonsense view of action and is rather obvious.

It may so happen that many different causal processes are capable of bringing about the same state or situation. If this is so then they may all be considered as instantiations of the same action; they are all equivalent under a certain (abstract) description since they all effect or bring about the same end. Consider again the door closing example. This action may be effected in any of the following ways: a person may push it shut; an animal may push it shut; it may be pulled shut; blown shut by the wind; attracted shut with a magnet; knocked shut by a flying object and so on. All instantiate the action of shutting the door⁷. Thus *action* is a term similar to *triangle* or *teapot*. Each is a name given to an equivalence class of entities: ‘triangle’ is the name given to a class of objects which have the same shape but may in every other regard be completely different; ‘teapot’ is the name given to a class of objects which are used to brew and hold tea but may differ radically in other regards; an ‘action’ refers to a class of processes equivalent only in respect of the effect they have of bringing about a specific state of affairs.

Affordances It is undeniable that it is perception that enables us (and other ani-

⁷ The point being made here may explain the curious assertion made by Reed that people do not achieve goals by “moving their muscles, nor . . . by displacing their limbs and bodies” (Reed, 1988; pp. 49). He may be referring, obliquely, to the fact that the achieving of goals (i.e., the performance of actions) is not to be *identified* with movements of the body.

mals) to act upon the world; after all, if we could not perceive the world we would not be able to do anything with it, at least nothing intentional. Perception tells us, at the very least, what we are able (and unable) to do with and in the world⁸. Gibson considered that when the world is perceived it is perceived in terms of what the perceiver can do — his actions. For example, if some piece of the environment is capable of being grasped by a person, then part of what is to perceive this piece of environment is to perceive that it may be grasped — that it is graspable. The property of being graspable is a description of the environment in terms of the act of grasping: properties such as graspability were called *affordances* by Gibson (1977). The notion that what is perceived is the ‘world’ described in terms of the organism is quite acceptable to many proponents of a computational approach. For example Koenderink (1980) observed that, “[s]olid shape’ is not present in nature but is a mutual property of perceiver and environment.” (pp. 390). Koenderink’s attitude appears to be that this is a metaphysical position that characterises the philosophy of the computational approach just as it does in the ecological approach.

Since there can be little doubt that we perceive how we can act upon the world, there can be little doubt that we perceive its affordances. There can be small cause for controversy here. What could be considered controversial, however, is the view expressed in the following quote from a recent paper by Carello, Groszofsky, Reichel, Soloman & Turvey (1989): “[t]he behavioural possibilities of an environmental layout taken with reference to an animal’s action capabilities is what is perceived.” (p. 29). What is controversial here is the view that the world is perceived in terms of an animal’s *own* actions, and most other discussions of affordances (e.g., Carello et al., 1989; Gibson, 1977; Shaw & Turvey, 1981; Shaw, Turvey & Mace, 1982; Turvey et al., 1981) treat perception in the same way. This view cannot, however, be considered complete. The reason is that to hold that an animal perceives its world in terms of its own actions is to assume that perception is egocentric in a way that seems very uncharacteristic of mature human perception at least.

Adult humans appear to be able to perceive the world in terms, not only of their own actions, but also of the actions of other agents and processes. For example, in perceiving that I cannot lift an object (relative to my own lifting capabilities it does not afford lifting) I may be able to perceive that it could be lifted with the aid, for instance, of other people or by means of a crane. Similarly, I can perceive

⁸ It can also tell us a good deal more than this as is discussed later.

how the actions of other animals, other people and natural processes will affect me, other people, other animals or physical objects etc. Our ability to perceive what the world affords the actions of other animals or natural processes allows us to exploit these animals and processes as tools with which to supplement our own actions. It appears that this point although compatible with Gibson's thinking and that of his followers has been overlooked by them.

It is perhaps worth noting at this point that a description on terms of actions and affordances can be given of any physical interaction whatsoever. Consider as an example the interaction between water and salt. Water may be viewed as performing the act of dissolving the salt and the salt may be viewed as 'affording' solution by the water. In general, consider any physical system, S_1 , which brings about some change, x , in a second system, S_2 , when placed in interaction with it. S_1 may be seen as bringing about x on S_2 and S_2 may be seen as affording that action for S_1 . Equally, such an interaction may effect a change y in S_1 and hence S_2 can be seen as acting on S_1 . Depending on whether one is interested in the change x or the change y , S_1 may be looked upon as actor or acted upon respectively. The language of action and affordance may thus be seen as simply a way of describing interactions between systems which is perspective dependant.

From the above discussion it may be suggested that a simple extension of the notion of affordance based perception is required if it is to effectively characterise human perception. It seems that our perceptual systems are telling us far more about the world than simply how we may act upon it, we are also able to perceive how other animals and people may act upon it, how it acts upon us and upon itself in the form of natural processes and how man made mechanisms and devices may act upon it.

There remains an apparent difference between the Gibsonian approach and the computational approach which is related to what has just been said. The ecological approach characterises perception as the extraction of information whereas the computational approach to visual perception is often viewed as taking vision to be a process that accepts the images and image flow as input and delivers a detailed three dimensional representation of the world as output — a view that ecological psychologists find objectionable (Carello et al., 1989; Turvey, 1986). From the discussion that has been presented so far in this chapter it may be argued, however, that the two characterisations of (visual) perception are in fact identical, they

simply use different language to describe the same thing.

It will be remembered from §1.2 that a representation is a ‘system’ for making certain information explicit (usable). Thus the representation of the world that is supposed in the computational approach to be delivered by perceptual processing is a ‘system’ (perhaps some kind of distributed array of neural activity) which makes the information about the properties of the world that we see — the shapes, solidity, roughness, size, distance, and colour of objects, the identity of these objects and so on — explicit. In other words, the computational approach understands vision as supplying a representation in which information about the properties of the world which people see when they look at their environments is made explicit. The ecological approach argues that the properties of the world that people perceive when they look about them are affordances. If this is a reasonable position, and computational theorists such as Koenderink (1980) think that it is, then we can read ‘affordances’ for ‘perceived properties of the world’ in the above characterisation of computational vision. Thus, the computational approach with an appropriate metaphysical background understands vision as supplying a representation which makes information about the affordances of the environment explicit.

It is worth noting, however, that it may not be at all clear to what affordance any given perceived property of the environment corresponds — consider Koenderink’s example of solid shape. Such properties as this seem to be summaries of all sorts of affordances. They may well be mutual properties of perceiver and environment (the position taken here) but they do not seem to be the affordances of particular actions. We might conceive of such properties as “meta-affordances” descriptions of the world not in terms of individual actions but in terms of whole classes of actions, i.e., descriptions in terms of what is common to the affordances of this whole class of actions.

As an example of the kind of thing we are trying to get at consider the perception of distance. We need to perceive the distance something is away from us if we are to reach over and pick it up, hit it with something, or throw something at it. If we are to do these things we need to perceive distance in a metric which is not specific to any one of these actions or any particular implementation of one of them. The metric should be abstract, i.e., common to all the actions and their implementations. Should we then decide on a particular action and implementation, the abstract distance information can be appropriately scaled. Returning to prop-

erties of the environment like solid shape, instead of seeing an object with which and upon which many actions can be performed in terms of all the corresponding affordances at once (e.g., as squeezable, throwable, graspable, edible and so on) one sees it as having properties, properties which represent only what is common to the affordances of all these actions. One only sees it in terms of a particular action when one decides to actually perform this action. In this way actions can be considered to “condense” out properties and thus we may say (with Koenderink) that these are mutual properties of organism and environment.

In computational vision (Marr, 1982) the content of conscious visual experience is taken as the indicator of what information the visual system extracts from the stimulus. It is not how someone behaves that tells the computational theorist what information that person is using, rather it is what properties of the environment he or she is consciously aware of. On the basis of the forgoing discussion it may be concluded that the computational theorists are right to adopt this approach — conscious visual experience is perhaps the only way of knowing about many of the environmental properties a person perceives. Nevertheless, such a strategy will clearly not work for animals where the only access to the content of their perception is through observation of behaviour — through action. In addition there is more to perception than detecting affordances. We not only need to perceive *what* can be done but also *how* to do it effectively; we need to be able to control our behaviour such that it can adapt to changing conditions. We turn now to this issue.

Information for Control There is much evidence to support the notion that conscious experience does not reveal to a person how their behaviour is actually controlled. It used to be thought, for example, that in catching a ball the grasp was initiated by a person when the ball was felt contacting the hand. This is not actually what happens — the grasp begins before the ball reaches the hand and is more or less complete before enough time has elapsed for a touch activated grasp to have even been initiated (Alderson, Sully & Sully, 1974; Judge & Bradford, 1988). Similarly people are not aware that they normally use vision to control posture (see e.g., Lee & Lishman, 1975) or, as Lee, Lishman & Thompson (1982) have shown, that in the long- jump visual information is important in guiding foot placement immediately prior to take off (it used to be thought that a standard run up is the key to successful jumping). Other examples of the phenomenon that people are unaware of how they control behaviour could be provided and some of them will

be discussed in some detail in the next two chapters.

The computational approach as presented by Marr (1982) and many subsequent researchers tends to restrict itself to the problem of creating a representation which reveals the information which a person is actually consciously aware of. Gibson and those whose research he influenced have long stressed that to understand what information is used to control action one must analyze and investigate the action rather than look at conscious visual experience (e.g., Gibson, 1958; Gibson, Olum & Rosenblatt, 1955; Lee, 1980a; Lee & Aaronson, 1974; Lee et al, 1982; Lee, Young, Reddish, Lough & Clayton, 1983; Lishman & Lee, 1975; Todd, 1981; Warren, 1984; Warren, 1988; Warren, Young & Lee, 1986). A similar position has been expressed by many of those researching the behavioural skills of animals (e.g., Arbib, 1972, 1981, 1987; Ingle & Shook, 1985; Reichardt, 1986; Reichardt & Poggio, 1976). The perceptual control of action will be discussed in detail in the next chapter.

§1.4. CONCLUSIONS

In this chapter argument and interpretation have been presented in an attempt to establish the compatibility between the fundamental concepts and ideas upon which Gibson's "ecological" approach is based and those upon which the computational approach is based. It was shown that both approaches differ from traditional conceptions in psychology by insisting that given ecological constraint the stimulus can (and often does) carry sufficient information to support veridical perception of the environment. Given that this was the founding rationale of Gibson's approach to perception, it is odd that the ecological and computational approaches should so often be seen as promoting completely different characterisations of perception. It was argued that this opposition arises from Gibson's controversial attitude to perceptual mechanism: an attitude based on a rather vague metaphor of perceptual "resonance", "tuning" and information "pick-up" following from an ingenious analogy with the reception of radio stations. This conception of mechanism has not proved to be scientifically fruitful and has conspired to confuse rather than clarify perceptual theorising.

Much effort has been expended and much convoluted argument generated by researchers in the Gibsonian tradition attempting to justify the rejection of information processing, computation and representation which seems to follow from the radio reception analogy. However, this argument has never found wide accep-

tance and a model of a perceptual system that follows the Gibsonian conception of mechanism has never been presented. It was argued here that a rejection of computational concepts in no way follows from the fact that the stimulus can be informationally rich and specific. Indeed, these concepts appear simply because one chooses to adopt a **functional attitude** towards system description — if one changes one's attitude towards the system these concepts become inappropriate modes of description. The rejection of computation, representation and so on as appropriate for the description of perceptual process may therefore be taken as tantamount to rejecting the functionality of perception. For this reason it seems most balanced to agree with computational theorists that computation and representation, appropriately defined, are necessary concepts for the description of perceptual process. This conclusion is independent of whether current computational models capture correctly the computational style of the brain. After all, the theory of computation is a young and developing field, especially now with developments in parallel processing and concurrency.

It may be thought that Gibson's insistence of the intimate relationship between perceiving and acting represents a difference between his approach and that of other theorists. However, it is difficult to justify this position. It was argued here that Gibson's notion of affordance has frequently been presented in a narrow and incomplete fashion. Extensions of the concept, in no way inconsistent with Gibson's thinking, appear necessary if human perceptual ability is to be fully captured. Once these extensions are made differences between the ecological and computational approaches become difficult to see. In addition, although it is true that Marr's work and that of many other computational researchers does not explicitly deal with the role of perception in the control of action, there are computational style theorists who have dealt with it (e.g., Arbib, 1972, 1981, 1987) and in the next chapter an explicit computational approach to perceptuo-motor control is presented.

It may be concluded, therefore, that it is likely to be a futile exercise trying to establish the directness or otherwise of perception and to argue about how best to study human perception — through action or through experience: the two are complementary. It is perhaps best to move away from vague theoretical and philosophical argument and concentrate instead on trying to rigorously formulate the problems of perception and perceptual control. The chapters which follow attempt a few steps along the path towards such a formulation for the case of

the perceptual control of timing and the nature of the perceptual information it involves.

Chapter 2 presents a general discussion of perceptual control relating it to control theory and theories of perceptuo-motor control in animals. This chapter is not essential for understanding the arguments presented in later chapters and hence can be omitted without loss to the development of these arguments. Chapters 3 to 7 are relatively self contained and deal with the question of what perceptual information is used to control the timing of actions involving the interception of moving objects. Chapter 3 briefly considers timing constraints in tasks such as catching and hitting which involve interception of a moving target and argues for the importance of time-to-contact information in achieving successful interceptions. Chapters 4 to 6 analyse what perceptual information about time-to-contact people might actually use to time interceptive actions and chapter 7 presents a series of experiments which test some of the results of this analysis. Chapter 8 presents a detailed summary of the arguments and results presented in earlier chapters and places them in a wider context.

2

Perceptual Control of Behaviour**§2.1. INTRODUCTION**

The study of human and animal motor behaviour used to be approached in biology and psychology from a data driven perspective — there were data to be accounted for rather than problems to be solved. Consequently, the fundamental problems of controlling a biological system such as the human body so as to perform useful tasks were not addressed in the literature for a long time. One of the first researchers to identify such problems and to point out their importance in understanding the organisation and function of biological motor control systems was Bernstein (e.g., Bernstein, 1967). The stir that Bernstein's work caused in the psychology of motor behaviour was akin to that produced by the work of Marr and Gibson in visual perception. The latter researchers were both interested in understanding the natural functions of perception rather than explaining bodies of data or addressing traditional theoretical questions. Since Gibson's approach stressed the function of perception in relation to action, it is perhaps not surprising that the affinities between Bernstein's approach to motor control and Gibson's approach to perception were quickly noticed and a union of the two forged (e.g., Kugler, Kelso & Turvey, 1980; Reed, 1982; Turvey, 1977; Turvey, Shaw & Mace, 1978). This new "theory of action", following Gibson, tended to reject computational concepts (e.g., motor programs) and notions from control theory (e.g., commands, signals, feedback) that were typically used in explanations of motor behaviour (Reed, 1982, 1984; Kugler et al., 1980; Soloman, 1988; Turvey et al., 1978). A "dynamical" approach was developed, purporting to be an alternative to an approach based on traditional computational and control theoretic notions (Kelso, 1981, 1986; Kelso & Schöner, 1988; Kugler et al., 1980; Kugler & Turvey, 1987).

A deep rhetorical divide has arisen separating the dynamical approach from traditional approaches and it is frequently implied that only one approach can be the "right" approach. In this chapter concepts from modern control theory are used to show that the two approaches do not, in fact, stand opposed but simply represent differences of emphasis.

§2.2. CONTROL

We talk of having control over something (a system of some kind: mechanical, electrical, economic etc.) when we can make it behave as we want it to behave. There will typically be certain things which we can do to the system which will cause changes to occur in the behaviour that we wish to control. By doing these things in an appropriate way the desired behaviour may be achieved (provided, of course, that it can be achieved in principle — the system is controllable). For example, we might want to make an automobile travel in a particular direction. The automobile's direction of travel can be changed by turning the steering wheel. Thus, the desired behaviour is to be achieved by turning this wheel appropriately. In order to do this it is clearly necessary to know how exactly to turn the wheel to achieve a particular change of direction. This example illustrates the three principle components of any control problem¹:

- A system of some kind that is to be controlled (usually called the **plant**).
- A specified pattern of behaviour that we require the plant to exhibit.
- Certain “things”—variously called **inputs**, **forcings** or **controls**—that can be done to the plant which act to change (in a predictable way) the relevant aspects of its behaviour.

The control problem itself is thus to devise a **scheme** or **method** which specifies how the available controls should be applied to the plant so as to cause it to display the desired pattern of behaviour. As an example, consider a commonly discussed problem in control theory texts (e.g., Barnett & Cameron, 1985), that of balancing an inverted pendulum. Corresponding to the three items described above we have:

- A “pendulum” attached to a shaft and free to rotate as illustrated in figure 2.1.
- The requirement that the pendulum be kept in the upright position for a “significant” length of time.
- The ability to exert torques on the shaft by means of the motor (see figure).

The control problem is to find a scheme for applying torques to the shaft so as to keep the pendulum balanced in the upright position. Without control the pendulum

¹ Other components need to be added to specify certain types of control problem such as optimal control problems or adaptive control problems.

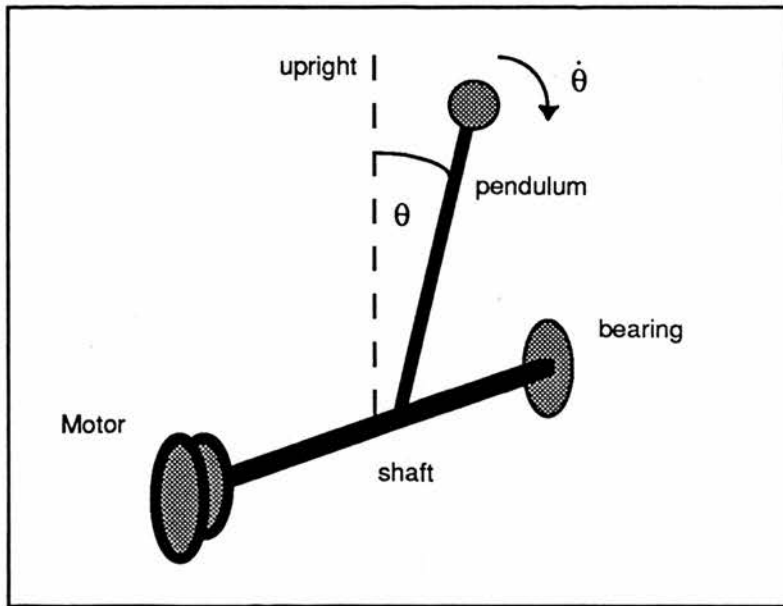


Figure 2.1. Pendulum fixed to a shaft which can be rotated by the motor. The pendulum is displaced from the upright by an angle θ .

is unstable in this position—it will tend to topple over and end up hanging vertically downwards (its equilibrium position, see below).

Schemes for applying inputs to a plant so as to achieve desired behaviour fall into two broad classes usually called **open loop** and **closed loop** schemes. In open-loop control schemes inputs are applied to the plant according to a pre-established schedule (a program) and in no way depend upon how the plant is actually behaving. This kind of control scheme is only appropriate when all the information relevant for achieving the desired behaviour is available at the outset, i.e., when no unforeseen events occur during the period over which control is exercised. Although open loop control is appropriate for some applications such as automatic toasters and washing machines, in general there will be unpredictable disturbances which may seriously compromise the effectiveness of such a control scheme. Under these conditions closed loop control schemes are appropriate. In a closed loop control scheme inputs are applied to the plant on the basis of how the plant is actually behaving. Before proceeding to give a precise formulation of the notion of closed loop control some important background material will be presented.

Control requires that the plant be given an appropriate description in terms of which the desired behaviour can be represented. It is meaningless to talk about

control without such a description. To make this statement more concrete, consider the example of a physical object described as an electrical system. In giving this description only the electrical properties of the object will typically be relevant and the description will contain only these properties. Other properties, such as mechanical or thermal properties, are typically ignored in an electrical description. The same physical object can, of course, be given a mechanical or a thermal description, but these descriptions will be independent of one another for most practical applications. Thus, if one wanted to control the electrical behaviour of a system, an electrical description is appropriate since it allows a representation the electrical behaviour that the system is required to exhibit².

In order to formulate a control problem and to devise a control scheme which solves it, a model of the plant must be provided in terms of which all its relevant behaviours can be represented. In the example of the inverted pendulum the physical system will typically be described as a purely mechanical object which can be completely described by giving the values of two variable quantities; the angle made by the pendulum with the vertical (θ) and the rate of change of this angle ($\dot{\theta}$)—the angular velocity. The mathematical model of a mechanical system falls into a class of models which are used universally in science as mathematical descriptions a physical system under study. These models are called *dynamical systems*.

A dynamical system³ is a mathematical object which relates the vector, $\mathbf{X} = (x_1, x_2, \dots, x_n)$, called the *state* to its rate of change with time, $\dot{\mathbf{X}} = (\dot{x}_1, \dots, \dot{x}_n)$, in the following way (which is also written out in component form for clarity):

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{F}(\mathbf{X}), \\ \frac{dx_i}{dt} &= f_i(x_1, \dots, x_n), \quad \text{for } i = 1, 2, \dots, n. \end{aligned} \tag{2.1}$$

Equation 2.1 says that the rate of change of state is a function only of the state. Real physical systems are modeled by dynamical systems of the form of 2.1 by representing what the physical system is like at an instant of time by a number of

² It is, of course, true that in some cases the electrical behaviour of a system is influenced by thermal or mechanical variables and so is, at least to some extent, potentially controllable using them. In this case the relevant variables become part of the electrical description. Ultimately one might suppose that all descriptions are parts of a single, unified description, but for practical purposes they can be considered independent as engineering practice and the history of physics attest.

³ The definition that is given here is not completely general but will be sufficient for illustrative purposes.

(potentially) observable physical quantities called **state variables**. The vector of values assumed by these magnitudes at that instant of time constitutes the system's state at that instant. For example, the state of a mechanical particle is completely specified by its position and its momentum which are thus the state variables of the particle. As mentioned above, the state can be viewed as a vector and thus the set of all possible states of the system constitutes a vector space having the same number of dimensions as there are state variables. This space is called the *state space*. Behaviour of the system is represented as a time-parameterised curve (trajectory) in this space. Thus, if we model the plant as a dynamical system, the behaviour that we desire it to exhibit may be represented as a particular trajectory in the state space. Returning to the inverted pendulum, the state of the pendulum is completely described by the observable quantities θ and $\dot{\theta}$ which are thus the state variables of the system. The required behaviour of the pendulum can then be represented as the "trajectory" in the pendulum's state space consisting of the single point $(\theta, \dot{\theta}) = (0, 0)$ (the "upright").

Of central importance in the study of dynamical systems is the notion of the *stability of critical points*. To understand this consider again equation 2.1 describing a generic dynamical system. This equation represents an assignment which associates with every possible state X (every "point" in the state space) a vector \dot{X} . An assignment like this which associates a vector with every point in a space is called a **vector field**. As a mathematical object, a dynamical system simply describes how every state in the state space is changing at any instant of time. In a physical system which can be modeled as a dynamical system, any change in state is considered to be caused by forces (force is simply a general term for any "agency" which causes changes in state⁴). A dynamical system of the form of equation 2.1, when used to model a real system may thus be considered to arise through the presence of forces internal to that real system which bind it together into a coherent object: in other words, *the rate of change of a system state is due to forces which depend solely on that state*. The *critical points* of a dynamical system are those points in the state space (i.e., states) at which there is no net "force"

⁴ In the case of a mechanical system where the state variables are position and momentum, this statement must be qualified since a non-zero momentum implies that the position is changing but a force need not be acting. This apparent contradiction is resolved when one considers that non-zero constant momentum in one coordinate system can always be considered to be zero momentum in another coordinate system. The change of state of a mechanical system is always defined relative to a special kind of reference frame called an *inertial frame*.

acting, i.e., points where the vector field vanishes. It would take us far beyond the purpose of the present discussion to provide a comprehensive discussion of stability, especially since our immediate purposes will be adequately served by considering just two simple concepts, **unstable states** and **asymptotically stable states**.

These two concepts are most simply illustrated by considering a well known and extremely simple model — a landscape of hills and valleys up and down which a ball can roll. If there is friction between the ball and the surface of the landscape then if the ball is rolled down a slope into a valley it will eventually come to rest in the very bottom of the valley⁵, and if the ball is then pushed a little way up the slope it will roll back down and come to rest in the bottom again. On the other hand, if the ball is placed on top of a hill it will remain there until pushed. However, if the hill top is not anywhere flat then the smallest push will send it down into a valley. The valley bottom can be considered as representing an asymptotically stable state and the hill top an unstable state. Basically, if a system is in an asymptotically stable state it will (eventually) return to it after a (small) perturbation, whereas it will not return to an unstable state.

To illustrate these concepts further consider again the inverted pendulum. The pendulum is in an asymptotically stable state when it is hanging vertically downwards and in an unstable state when poised in the upright position — the pendulum will return to the hanging down position when pushed but when it is pushed while poised in the upright position it will not return to the upright. There are forces acting on the pendulum which act to make the hanging down position stable and the upright position unstable.

Notice that nothing whatsoever has been said about the *mechanism* of the forces acting. This is a characteristic of dynamical modelling — it is completely mute as regards mechanism. This brings us back to the notion of dynamical analogy introduced in chapter 1 — the same dynamical system (a mathematical object) may be used to describe different types of physical system involving completely different physical mechanisms. This sort of thing is fundamental in physics: Feynman, for example, remarks that, "...there is a most remarkable coincidence: *the equations for many different physical situations have exactly the same appearance.*" (Feynman, Leighton & Sands, 1964; p. 12-1, italics in original). Feynman goes

⁵ Provided, of course, that it has not gathered enough momentum to enable it to climb up the other slope and over the top into the next valley.

on to give some examples and notes that, for instance, “*steady heat-flow problems and electrostatic problems are the same*” (*op. cit.* p. 12–2; italics in original). These two kinds of mechanistically quite distinct physical situations are dynamical analogs⁶.

It follows from the above discussion that the essential problem of control is one of transforming the dynamical description of the plant (the plant is, after all, a dynamical system) into the dynamical description of a system which exhibits the desired (controlled) behaviour. That is, the coupled plant–controller system (the control system) must be describable as a dynamical system which exhibits exactly the behaviour that we require. The control problem is thus the problem of finding a controller which does just that. These points will be discussed in more detail in the next section where the notion of a closed loop control scheme is stated precisely.

§2.3. CLOSED LOOP CONTROL

Continuing with our consideration of the pendulum balancing example, observe that the pendulum’s natural stable equilibrium position (i.e., hanging down) has precisely the stability properties that we desire of the upright position. We require that the upright be (asymptotically) stable in the face of perturbations. In other words, the problem of balancing an inverted pendulum can be stated to be the problem of turning the upright position from an unstable state to an asymptotically stable state.

The “hanging vertically down” position is a stable state of the uncontrolled pendulum by virtue of the forces that act naturally on it. To give the upright position similar stability properties we need to supply additional forces to the system. As stated above in our specification of the pendulum balancing control problem, we are able to supply forces (inputs) to the pendulum by means of the torque motor (see figure 2.1). Our problem, of course, is to supply the *right* forces — this is the role of the control scheme or, as it is more usually called in the context

⁶ The notion of analogy pervades the whole of physics one supreme example being Hamilton’s exploitation of an analogy between particle mechanics and geometric optics which led to the formulation of the Principle of Least Action and the Hamilton-Jacobi equation (Lanczos, 1970). In recent years many remarkable dynamical analogies between systems as diverse as populations of interacting biological species, magnetism in extended media, chemical reactions, networks of automata (neural networks), and cell differentiation during development have been discovered (see e.g., Amit, 1989; Haken, 1983; Rosen, 1985). It is this that forms the basis of the field Haken calls *synergetics* (Haken, 1983).

of closed loop control, the *control law* (or sometimes, the control algorithm).

The forces which give the hanging downwards position of the uncontrolled pendulum its stability properties depend only on the *state* of the pendulum. In a like fashion, the forces which are applied to the pendulum by the torque motor should be functions of the pendulum state — the forces used to control the pendulum should be “computed” from the state. At this point one can do no better than to quote from the founder of modern control theory, R. E. Kalman:

The principle that the inputs should be computed from the state was enunciated and emphasized by Richard Bellman in the mid-1950's. *This is the fundamental idea of control theory.* (Kalman et al., 1969; italics in original).

In accordance with this principle the following definition of a control law can be given (see Arbib, 1973; Kalman et al., 1969):

Definition. A control law is a map $\kappa : X \rightarrow U$ that assigns to the state $\mathbf{X}(t) \in X$ at time t the value $u(t) = \kappa(\mathbf{X}(t))$ as the value of the input to the plant at that time.

In this definition, X is the set of possible states of the plant (the state space) and U is the set of admissible inputs. Thus a control law just specifies what input should be applied to the plant when it is in a given state. This definition of control law formalises the notion of *feedback control*. The physical realisation of a control law is a device known as a controller. It is clear that in order to implement a control law the controller must be able to do two things:

- It must be coupled to the plant via a mechanism which applies the inputs.
- It must have access to information about the state of the plant in order that it may compute the appropriate inputs — it must have a way of *observing* the plant.

The set up that one ends up with is illustrated schematically in figure 2.2 and is called a *control system*.

What kind of system is the control system formed by coupling a plant and a controller? Consider the pendulum example, suppose that the uncontrolled pendulum system (the plant) can be described as a dynamical system of the form given

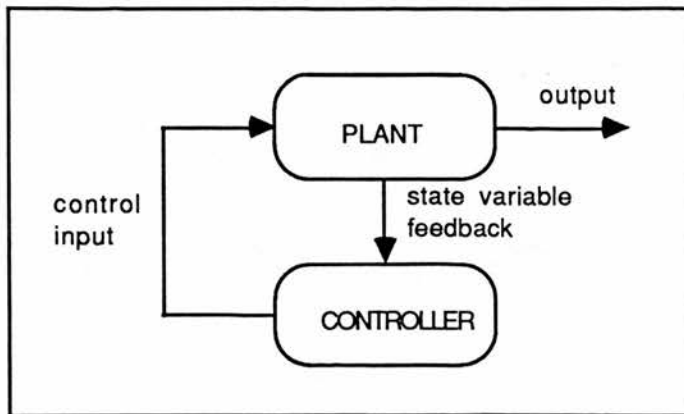


Figure 2.2 Simple feedback control system.

by equation 2.1. In this way, calling the angle θ , θ_1 , we may write,

$$\begin{aligned}\dot{\theta}_1 &= \theta_2, \\ \dot{\theta}_2 &= f(\theta_1, \theta_2).\end{aligned}$$

In this equation the state of the pendulum is represented as (θ_1, θ_2) . If a controller is coupled to this system and implements a control law which applies the input $\kappa(\theta_1, \theta_2)$ to the plant when it is in the state (θ_1, θ_2) we have, instead of equation 2.2 the following expression,

$$\begin{aligned}\dot{\theta}_1 &= \theta_2, \\ \dot{\theta}_2 &= f(\theta_1, \theta_2) + \kappa(\theta_1, \theta_2) \\ &= g(\theta_1, \theta_2).\end{aligned}$$

Which is clearly just another dynamical system which now includes forces which are due to the controller. As Hogan has pointed out,

It is impossible to devise a controller which will cause a physical system to present an apparent behaviour to its environment which is distinguishable from that of a purely physical system. (Hogan, 1985, pp. 1)

Thus a control system (figure 2.2) may be modeled as a dynamical system just like any other physical system, as is clear from any text dealing with modern control theory (e.g., Anand, 1984; Barnett & Cameron, 1984; Kalman et al., 1969; Luenberger, 1979). Indeed, it is possible to establish mathematically the general equivalence between the stability of steady states in dynamical systems and negative feedback control (see e.g., Rosen, 1985): in the neighbourhood of a steady

state a dynamical system “may itself be decomposed into a part interpretable as a controlled system [the plant in our terms], and a part interpretable as a feedback controller” (Rosen, 1985, pp. 40).

An important point that should not be overlooked is that a control law is, like the concept of a dynamical system, mute with respect to the actual physical mechanism which implements it. Figure 2.2 represents the *functional* organization of the control system but not necessarily its *mechanistic* organisation. It can be considered to represent *information flow* within the control system. It is tempting to imagine that a controller must consist of the following three physically identifiable things:

- (1) A measuring instrument or sensor which measures the output of the plant.
- (2) A controller which uses the information provided by the sensor to compute the appropriate inputs (and perhaps an observer to compute the state of the plant if the output data alone is insufficient).
- (3) An actuator which, under instruction from the controller, applies the computed inputs to the plant.

Whilst it is often convenient to think of a control system in these terms it can sometimes be a trifle misleading. It tends to make one think of measurements being made and transmitted as signals to a controlling device which then transmits signals (commands) to the actuator. Although this may sometimes correspond to what is actually going on it is by no means always the case. Consider the classic example of a feedback control system, Watt’s centrifugal engine governor. A schematic view of a steam turbine equipped with a governor is given in figure 2.3a and its operation is explained in the caption.

The engine governor can be represented in terms of a diagram of the kind presented in figure 2.2. Such a diagram is shown in figure 2.3b. Although the system can be represented by this diagram, it would be misleading to suppose that the operations of sensing the turbine speed, finding the appropriate opening or closing of the steam valve, and actually opening or closing this valve were physically separable and identifiable events. The mechanism of control in this case does not consist of a measuring device sending signals to a controller that then sends commands to the actuator. What we have is a mechanical device consisting of the flyballs connected to a system of rods which move the valve appropriately.

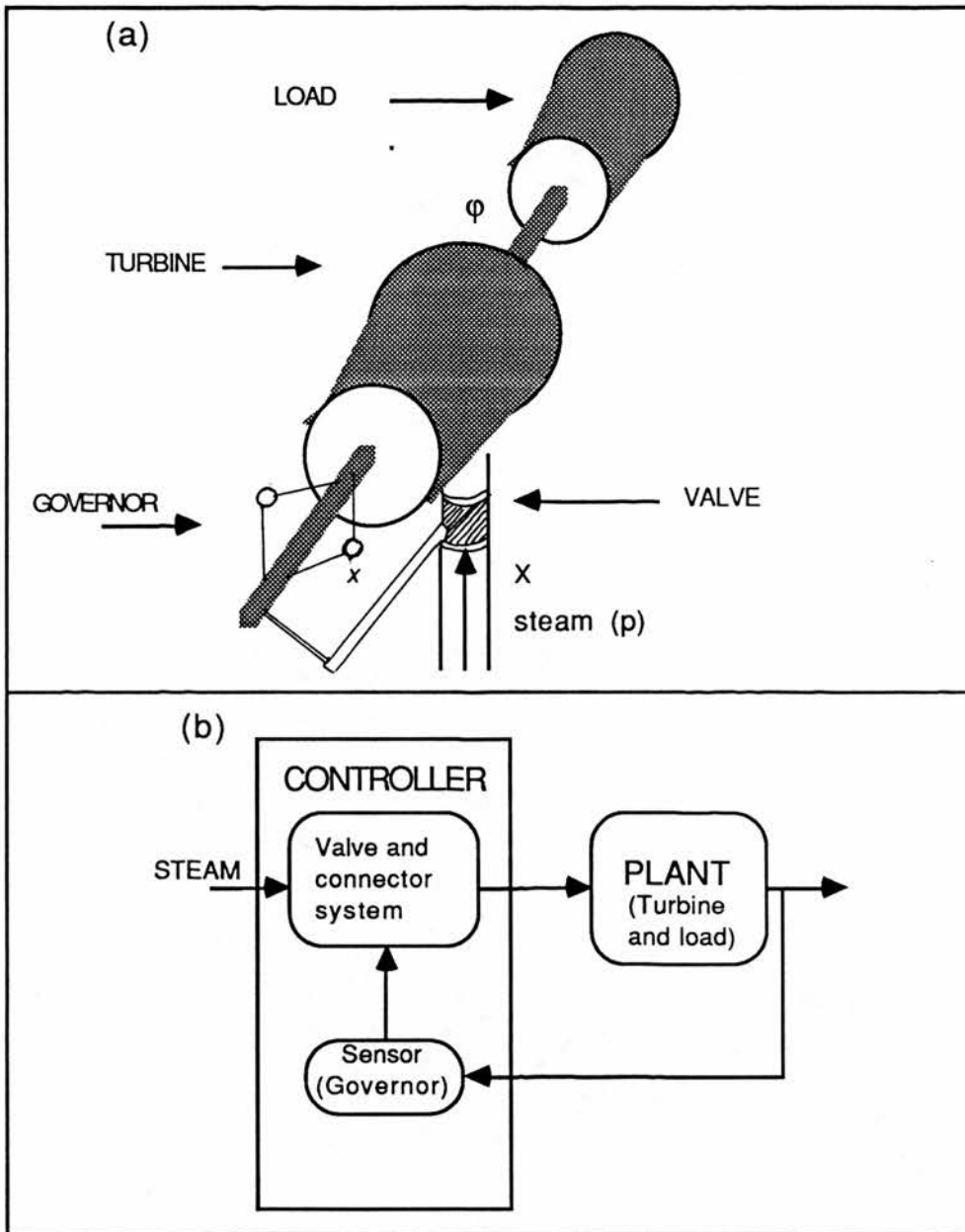


Figure 2.3. a) Perspective view of a steam turbine engine, governor and load after Olson (1958; figure 13.4, pp. 216). p is the input steam pressure; X is the volume velocity of the shaft; x is the velocity of the valve and ϕ is the angular velocity of the turbine shaft. b) Control system block diagram of the engine-governor system showing "information flow" within the system.

This device is *functionally equivalent* to a device which explicitly measures the turbine speed sends signals to a controller which then commands the actuator (valve). In other words, the diagrams given in figures 2.2 and 2.3b should be conceived of as *functional descriptions* of a control system not necessarily as explicit

descriptions of mechanism. It will be remembered from chapter 1 that it is adopting a functional attitude towards system description that makes the term ‘computation’ an appropriate term for certain physical processes. In the case of control it is the adoption of a functional attitude that makes talk of signals, measurement and commands appropriate, they are functional terms.

These examples illustrate the point that when a functional perspective is adopted it is appropriate to talk of information flow within a system and to use diagrams such as figures 2.2 and 2.3b. This kind of description helps to us to understand how a system functions. But we must remember that a functional decomposition of a system into functional units like sensor, controller, comparator and so forth does not necessarily imply a mechanistic decomposition into physically separate, identifiable functional components. Function may be distributed over the physical “parts” of a system such that the same part is involved in the implementation of perhaps several different functions simultaneously. In short, to describe a physical system as a control system is to take a functional perspective. It then becomes meaningful, and useful, to describe the system as having a functional organisation such as that illustrated in figure 2.2 or 2.3b. It is meaningless to argue over whether or not the system ‘really’ has such an organisation since the organisation is a product of the perspective that is taken towards system description.

One feature of many practical control systems, such as those described above, is that they are *tunable*. That is, certain parameters of the controller can be adjusted so that, for example, the engine stabilised by the governor can be controlled so that it runs at any speed one may require (within certain working limits). In control system block diagrams the ability to modify the behaviour of the system is represented by an arrow entering the controller labelled “command input” or “reference signal” (figure 2.4).

As stated above, the control system in figure 2.4 can be modeled as a whole as a dynamical system. The control system converges to a state (or state trajectory) which represents the behaviour that we require from it. The idea of the command input is that if we require different behaviour from the control system at different times, then by changing the command input appropriately we can change the dynamical behaviour of the control system such that it converges to a different stable state corresponding to the new required behaviour. For example, if we required

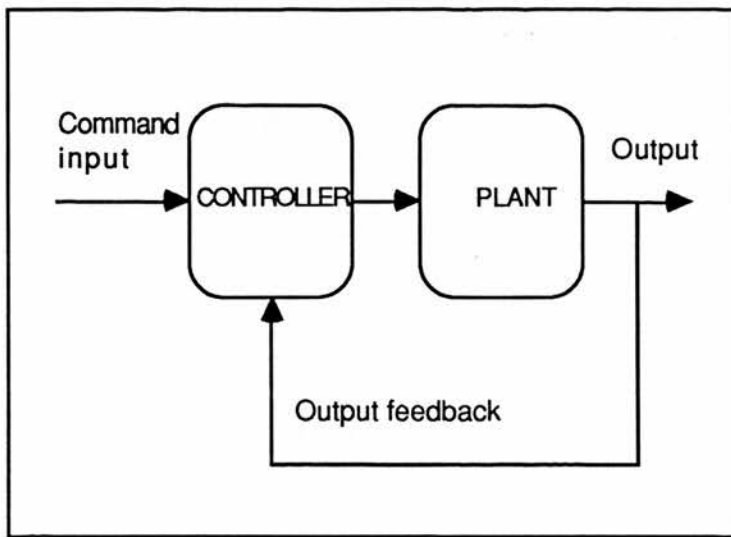


Figure 2.4. Simple feedback control system which follows an external "command" input.

the engine to run at a different speed we could adjust the governor in such a way as to stabilise running at this new speed. In effect, the control system in figure 2.4 is a tunable dynamical system. The command input is actually controlling the behaviour of the control system itself and so what we have now is open loop control of a plant which itself is a control system. This can be seen pictorially by extending the arrow labelled command input back and adding in the source of this input as in figure 2.5a. Clearly, it is possible to imagine a nested hierarchy of control systems and this is suggested in figure 2.5b.

In this section the notion of closed loop or feedback control was described. The following points were made:

- A feedback control system can be described as a dynamical system in the same way as any other physical system.
- One of the major roles of negative feedback is to make a desired state trajectory of the plant stable. In the case of the pendulum example, feedback control transforms an unstable state of the plant into a stable state.
- A control system may be functionally decomposed into a sensor (and if necessary a state observer), a controller and the plant itself. However, this is a

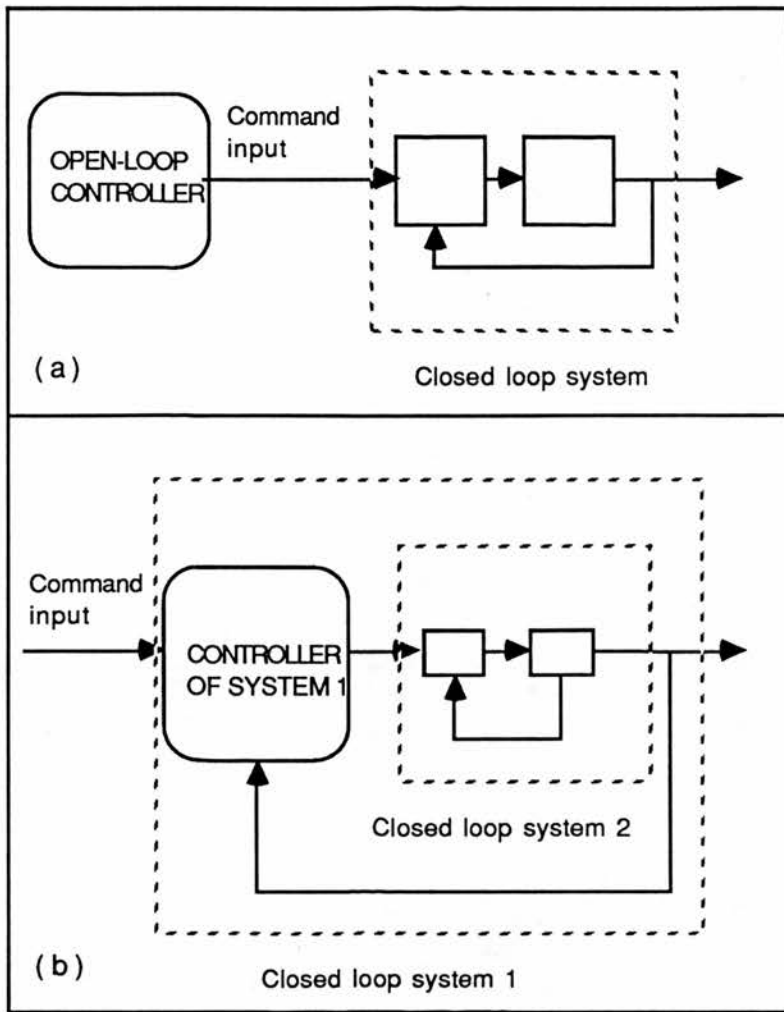


Figure 2.5. a) Closed loop control system as the system controlled by an open-loop controller. b) Open loop control of a closed loop system (closed loop system 1). The plant of the system under open loop control is itself a closed loop system.

functional decomposition and in no way implies that there be *physically* separate mechanisms which correspond to the *functionally* separate components.

- One can suppose that the plant of a control system is itself a control system and by this means build up hierarchies of control systems.

§2.4. PERCEPTUAL CONTROL: FIRST EXAMPLE

As discussed above, a control system may be conceived of as a dynamical system in the same way as any physical system can be so considered. The design of a control system involves finding a control law which leads to the system having the required dynamical behaviour (in some specified operating range). In these terms control system design is the creation of a dynamical system having desirable properties (Anand, 1984; Rosen, 1985). A behaving animal presents the opposite problem: we have a control system of unknown design which we wish (perhaps) to understand. The design problem suggests that to understand such a control system we need to have an overall description of its dynamics —we need to discover the dynamical system which describes the control system in its operating range. Only then can we begin to understand the mechanisms which implement the controlled behaviour, i.e., we have to know precisely what the controlled behaviour is before we can hope to understand it. This, of course, is exactly the problem that was described earlier in the context of perception. In chapter 1 it was argued that in order to understand a perceptual system a computational theory is required. Such a theory involves identifying a perceptual task — the perceiver needs to perceive such and such a quantity or property (e.g., colour, distance or direction of locomotion) — and understanding the information which the perceiver uses to obtain the perceived quantity. Before we can hope to understand a perceptual system we must understand precisely what it is doing, and this is conceived generally as picking-up information to solve a perceptual task. In the context of investigating a control system of unknown design, a computational type of theory is required which specifies exactly what the controlled behaviour is that we wish to understand. A detailed analysis of one kind of biological control system has been carried out which illustrates the type of approach that captures all that has so far been said concerning control and perception. It is the analysis of the flight orientation behaviour of the housefly (*musca domestica*) provided by the research of Reichardt and his colleagues (for reviews see Reichardt, 1986 or Reichardt & Poggio, 1976).

Reichardt considers the task of orienting to a visually specified environment or object in the environment. He observes that "... Male and female flies fixate—that is fly towards—small contrasted patterns and they track moving objects." (Reichardt, 1986, p. 113). The problem is to characterise and understand the control system which underlies this behaviour. Reichardt and colleagues have studied the

fly's orientation behaviour experimentally and have discovered that horizontal and vertical control are more or less independent. For ease of presentation only control in the horizontal plane will be considered here.

The analysis begins by describing the dynamics of the fly whilst it is engaged in the act of orienting. That is, the orienting fly is characterised as a dynamical system. The experimental investigations ignored the effects of the fly's translation through its environment and thus correspond to the situation in which the object(s) being oriented to are too far away for their position relative to the fly to change significantly with translation⁷. We are thus concerned with the rotational flight dynamics in the horizontal plane. It was found that the rotatory flight dynamic of the fly can be well approximated by

$$\Theta\ddot{\psi}(t) + \beta\dot{\psi}(t) = -F(t) + S(t). \quad (2.6)$$

Θ is the fly's moment of inertia, β an aerodynamic friction constant of the fly, and $F(t)$ is the instantaneous input torque provided by the fly's wings which thus represents the means by which the fly can control its flight. The angle ψ is defined in the following diagram (figure 2.6) and represents the error in fixation — the angle between the instantaneous direction of flight and the direction of the object which is to be fixated. The quantity $S(t)$ ($= \Theta\ddot{\alpha}_p(t) + \beta\dot{\alpha}_p(t)$) represents the motion of the object within the fixed frame of reference (if the object is stationary, $S(t) = 0$).

Extensive experimental investigations (reviewed by Reichardt & Poggio, 1976) establish that the input torque can be approximated to well within experimental error as the sum of three terms as follows:

$$F(t) \approx D(\psi(t)) + r\dot{\psi}(t) + N(t). \quad (2.7)$$

The last term, $N(t)$, is a gaussian process representing intrinsic noise in the control system. The first two terms represent how the fly's flight control system transduces visually detected error angle ψ and angular velocity $\dot{\psi}$ into torque exerted by the wings. Equation 2.7 describes the control law implemented by the fly (compare equation 2.7 with equation 2.3)—it is a mapping which assigns to each state of the fly as described by ψ and $\dot{\psi}$, a control torque F according to the rule given by equation 2.7. Neglecting the visuo-motor delay involved in transducing visual

⁷ Taking translation effects into account (Reichardt & Poggio, 1981) complicates but does not alter the picture in any essential way and so it is quite reasonable to omit them here for the sake of illustration.

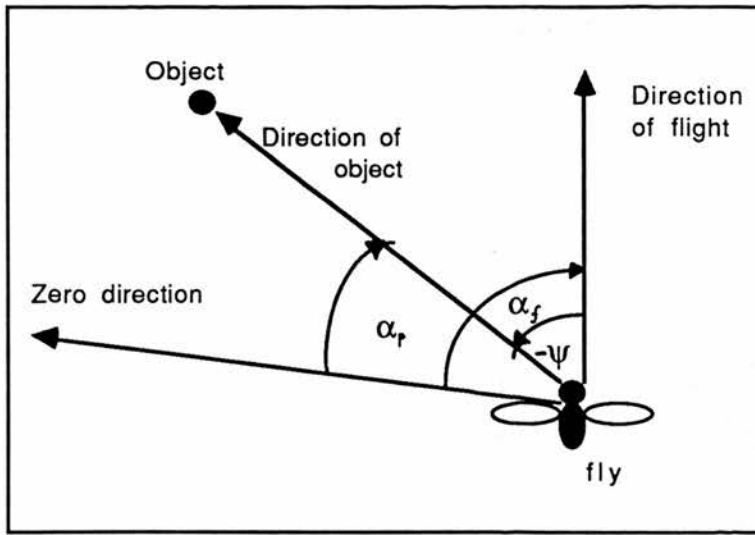


Figure 2.6. Coordinate system, after Reichardt & Poggio (1976). The object is effectively so far from the fly that the fly's translation does not affect the direction of the object. The angle α_p therefore only changes if the object moves relative to the fixed frame of reference. The "error" angle, ψ , is equal to $\alpha_f - \alpha_p$ and can change if the object moves or if the fly rotates (or both). When $\psi = 0$ the object is fixated.

information (ψ and $\dot{\psi}$) into torque response since it is very small (less than 20ms), equation 2.6 may, after substitution from equation 2.7, be rewritten as follows (Reichardt & Poggio, 1976):

$$\Theta \ddot{\psi}(t) + (\beta + r)\dot{\psi}(t) + D(\psi(t)) = N(t) + S(t). \quad (2.8)$$

The above equation describes the dynamics of the flight control system at the level at which its function is defined. To see this more clearly, consider fixation of a stationary object ($S(t) = 0$); taking the noise to be negligible equation (2.9) becomes

$$\Theta \ddot{\psi}(t) + (\beta + r)\dot{\psi}(t) + D(\psi(t)) = 0. \quad (2.9)$$

The experimentally determined form for the function $D(\psi)$ is illustrated in figure 2.7a and may be considered to be the derivative of the potential illustrated in figure 2.7b (Reichardt & Poggio, 1976). With $D(\psi)$ so defined the dynamical equations describing the fixation behaviour of the fly have an equilibrium position at $\psi = 0$ — the bottom of the potential well (figure 2.7b). The fly exerts torques which tend to reduce ψ to zero. Thus, the perceptuo-motor processes of the fly's

equations describing the fixation behaviour of the fly have an equilibrium position at $\psi = 0$ — the bottom of the potential well (figure 2.7b). The fly exerts torques which tend to reduce ψ to zero. Thus, the perceptuo-motor processes of the fly's flight control system give the fly the dynamical behaviour captured by equations 2.8 and in the case of noise free fixation of a stationary object, equation 2.9.

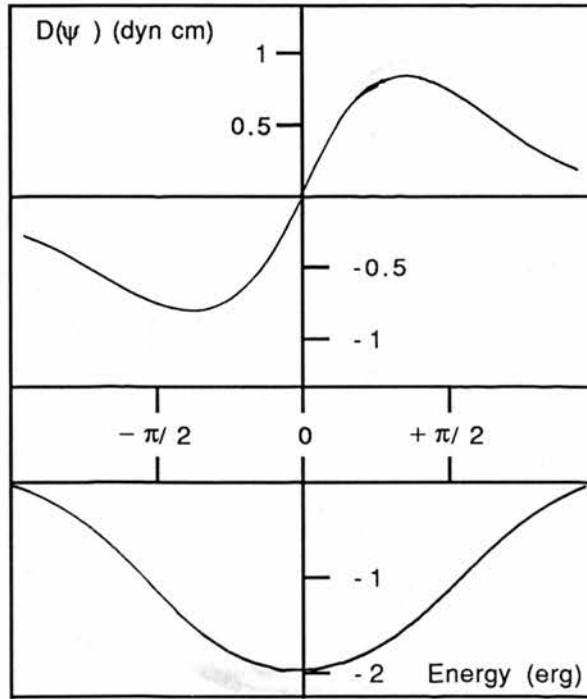


Figure 2.7. a) The form of the empirically determined function $D(\psi)$. b) The potential of which $D(\psi)$ in (a) is the derivative (after Reichardt & Poggio, 1976, figure 7).

From the point of view of this purely behavioural analysis, relevant fly behaviour is completely described in terms of a dynamical system where the control torque components D (figure 2.7a) and r may be considered as rotational “stiffness” and “friction” (damping) respectively. The fly behaves as if its flight direction were attached to the direction of the object by a (slightly fluctuating—noisy) spring.

Observe that no subtraction from a set-point is required of the fly's control system in Reichardt's experimental set-up (Reichardt & Wenking, 1969) since the direction of flight corresponds to the visual “straight ahead”. The angle ψ is thus just the angle of the object measured relative to the straight ahead and can presumably be measured directly. The fly's task is essentially one of keeping the

set point (desired equilibrium position) can be built into the device which measures the angle, i.e., the angle is measured from the “set point” directly, obviating the need for any subtraction.

It is clear from the above discussion that in Reichardt’s analysis the flight control system of the fly is characterised in the same way as the control strategy for balancing an inverted pendulum described in section 2. There is a dynamical system (plant) that is to be controlled so as to produce a desired behaviour; in the case of the flight control system this is the body of the fly in air. Without control the fly’s body has no preferred direction. When coupled to a controller the fly’s body does have a preferred direction — the direction of the fixated object. The fly with control now has a stable state where the angle ψ and its rate of change $\dot{\psi}$ are both zero, just as the inverted pendulum with control has a stable state (the upright). In both cases, the system with control is a dynamical system in which some of the forces are due to perceptual processes rather than to simple mechanical ones. The “rule” which maps perceptions onto applied forces is a control law. In the fly’s control system the control torques are simple functions of the perceived quantities (ψ and $\dot{\psi}$) just as in the pendulum balancing example the control torques are simple functions of the “perceived” quantities (θ and $\dot{\theta}$).

It is now appropriate to make the following observation. As pointed out in chapter 1, researchers in the Gibsonian tradition have long emphasized that perception and action are not to be considered as independent systems that can be treated as logically separate but rather are to be considered as two aspects of a unitary ‘perception-action’ system (von Hofsten, 1987; Lee, 1980b; Turvey, 1977; Turvey & Carello, 1986; Warren, 1984, 1988). A feedback control system is a perception-action system in the following sense: measured (perceived) quantities (and quantities estimated from the measured quantities) specify the controls (inputs) that are to be applied to the system. In both the fly and in the pole balancing system the exerted torques are determined by perceived quantities (the control law specifies the torque to apply given values of the perceived quantities). A control law binds together perception and action into a unitary (control) system:— perception specifies action and action (behaviour) determines perception. In a recent attempt at formulating Gibsonian intuitions about perception and action Warren (1988) was led to the notion of a feedback control law but was either unaware of this fact or neglected to mention it. Warren uses the term “control law” to mean exactly

what it means according to the definition given earlier. Since he does not reference the control literature in which this definition appears (see references given earlier) it would appear that he is unaware of it. The notion of a perception-action system as it has been employed in the ecological literature would seem to correspond to the notion of a control system. Given this equivalence it is possible to be more precise about the problems of perceptuo-motor control than has been typical in the existing psychological literature on the subject. In the next chapter an analysis of the control of timing is given which makes use of the concepts that are developed in this chapter.

In this section an example of a control systems analysis of a behaving biological system has been described. This analysis characterizes the flight orientation behaviour of the housefly as a dynamical system. The "goal" of the act of orienting is represented as a state in which the angle between the fly's line of flight and the direction of an object of interest (the error angle) is zero and unchanging (allowing for a bit of noise). The fly acts so as to bring its state (as described by the error angle, ψ , and its rate of change, $\dot{\psi}$) from wherever it may be initially to the goal state. This behaviour is completely described by a dynamical system defined on the space of states whose attractor is the goal state.

The formulation and analysis of the problem hinges on the characterisation of the task of orienting the direction of flight to a visible target as that of bringing ψ and $\dot{\psi}$ to zero⁸. This is an abstract characterisation which is independent of the means by which the orientation is to be achieved, it just says that orienting requires that an axis be pointed in a particular direction. In this sense the characterisation applies equally to orienting the fovea of an eye towards an interesting object, pointing a missile at a target or driving a car at a target. Indeed, as pointed out by Reichardt & Poggio (1976), the analysis they provide for the fly's orienting behaviour could be used to model pursuit eye movements in man — the tasks are abstractly equivalent. The abstract characterisation of orienting is very simple but it serves to illustrate an important point: tasks can be characterised independently of the system which performs them. This, of course, is just another way of saying that different systems can perform the same tasks; the tasks must therefore be, in some sense, independent of these different systems. The act of orienting to a target and the solution as implemented by the fly are shown schematically in the

⁸ Bearing in mind, of course, that the analysis was confined to the horizontal plane.

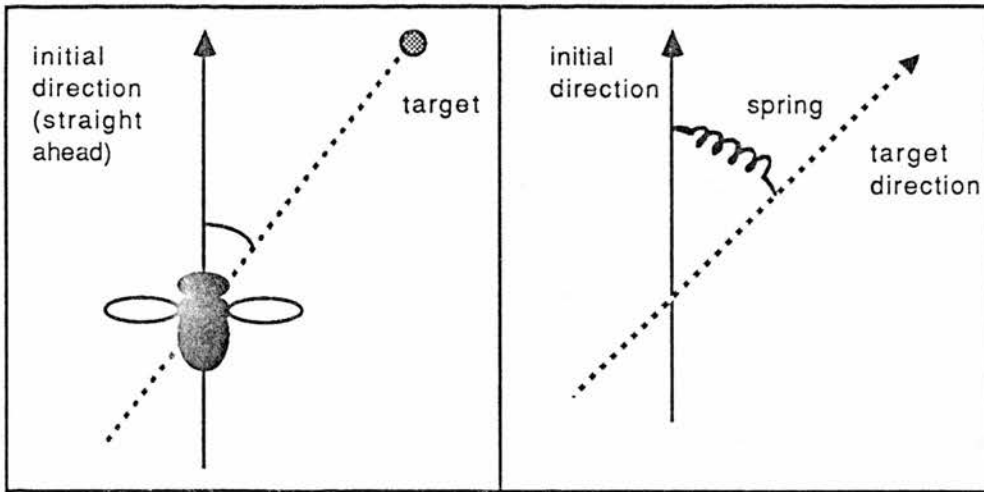


Figure 2.8. a) The orientation task of the fly is that of pointing a reference axis (the straight ahead or direction of flight) in the direction of the target. b) The solution of the orientation task as implemented by the fly is equivalent to linking the two directions by a "spring" which pulls on the straight ahead direction if it is not coincident with the target direction.

following diagram (figure 2.8).

It is quite routine to characterise tasks abstractly (i.e., independently of any particular device which might carry them out) in control engineering and (especially) robotics (see e.g., Brady et al., 1982). As is apparent from the above discussion such abstract characterisations correspond to the common sense view of tasks (see also the discussion of action presented in chapter 1). The corresponding abstract characterisation of the solutions to tasks as dynamical systems on state spaces defined at the task level, as was done for the pole balancing task and the fly's flight orientation, has not been done until recently in either robotics or in the study of biological motor control. This strategy will be discussed further in section 2.6 below.

§2.5. PERCEPTUAL CONTROL: SECOND EXAMPLE

The example of the last section described a situation in which behaviour was *continuously* controlled by visual feedback information corresponding to the control scheme of figure 2.2. In this section, control of movement according to the scheme illustrated in figure 2.5 is considered. An explicit example of such a scheme is Feldman's λ model of motor control (see e.g., Feldman, 1986). Figure 2.9 shows a rough and schematic neurophysiological interpretation of the mechanistic structure of the model (Berkinblit, Feldman & Fukson, 1986). The muscle plant together with the closed-loop control mechanisms behaves in a spring-like fashion with (static) exerted force depending on muscle length in a characteristic fashion (for a discussion see Feldman, 1986). The resting (equilibrium) length the muscle assumes when loaded depends on the load, the characteristics of the muscle and its associated closed-loop mechanisms, and the central commands.

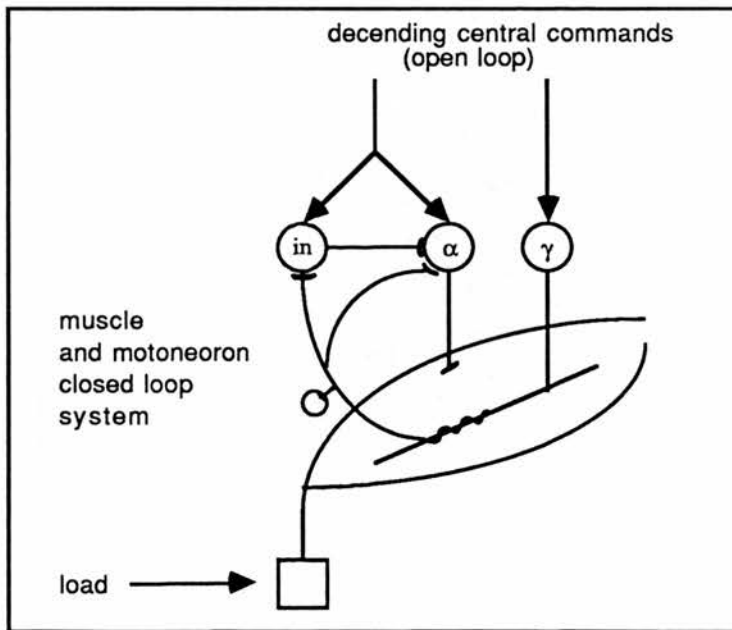


Figure 2.9. Neurophysiological interpretation of Feldman's λ model of muscle control (after Berkinblit, Feldman & Fukson, 1986, figure 4.). The muscle and motoneuron system form a closed-loop control system which allow the muscle to behave like a spring with a variable equilibrium length. The central commands are open-loop with respect to the behaviour of the muscle (but not necessarily with respect to the task which the animal is using its muscles to perform).

Feldman's model asserts that the muscle length at which feedback from the muscle leads to increased exerted force due to α motoneuronal activity can be varied in an open-loop fashion by central commands: this variable is the threshold, λ (hence λ -model), of the stretch reflex (Feldman, 1986). By changing this threshold the equilibrium length of the muscle changes and hence may be controlled by setting the threshold. This scheme is exactly analogous to thermostatic control of the temperature of an appliance like a fridge. The thermostat itself is a closed-loop device which tends to stabilise a particular temperature (the temperature when the thermostat reaches equilibrium). The actual temperature stabilised can be set open-loop, typically by turning a dial. Notice that setting such a dial to a particular value defines a new equilibrium state for the closed-loop system and so setting it to one value and then to another value and then to another and so on at subsequent (equal) intervals of time (Δt) defines a sequence of equilibrium states. In the limit ($\Delta t \rightarrow 0$) we have a continuously changing setting on the dial which one can think of as defining a continuously changing sequence of equilibrium states for the closed-loop system which tracks them providing that the response of the closed-loop system is sufficiently fast. In the context of muscle control such a sequence of equilibrium states has been referred to as an "equilibrium point trajectory" (Bizzi, Accornero, Chapple & Hogan, 1984).

Consider now a simple single joint positioning task which has been used by Bizzi and co-workers to study the spring-like properties of deafferented muscle control in primates (Bizzi, 1980; Bizzi et al., 1982, 1984), though afferented muscle will be considered here. The task involves extending the elbow joint so that the hand points to a visual target light. There is no other illumination and so the pointing is done without visual feedback from the limb itself. The limb comes to rest at a particular joint angle (φ) when the torque exerted by the elbow flexors is exactly balanced by the torque exerted by the extensors: this is the equilibrium posture of the limb. The two sets of muscles act together in determining the equilibrium joint angle as an agonist-antagonist pair. Feldman (1980a,b) identifies two forms of open-loop central command which can control the pair of muscle groups influencing joint angle. The first of these, called "coactivative", sets the stiffness of the joint and is not of direct concern to the present discussion. The second type of central command, which Feldman terms "reciprocal", is the means by which the equilibrium angle can be set. The logic of the operation of this kind

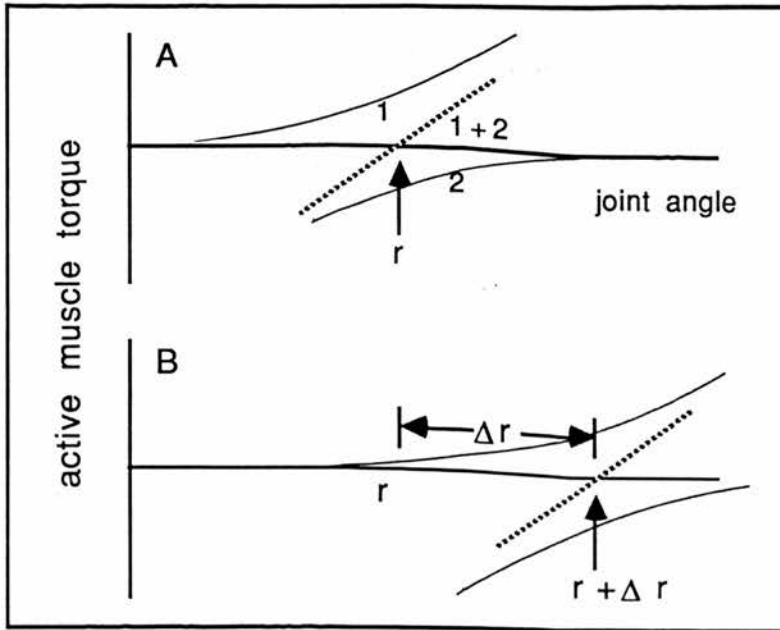


Figure 2.10. Definition of reciprocal central motor commands in terms of shifts in the torque angle functions (invariant characteristics or ICs) of flexor (1) and extensor (2) muscles at a joint (after Feldman & Latash, 1982, figure 2). The IC of the combined muscles is shown as a dotted line (1+2). The reciprocal command shifts the IC of each muscle by the same amount (Δr) and thus the overall IC is shifted from r (A) to $r + \Delta r$ (B).

of control signal is illustrated in figure 2.10.

The reciprocal command controls the agonist and antagonist muscle groups as a unit whose joint action determines the equilibrium angle of the limb (given a specified load). This unit itself is capable of acting autonomously in a closed-loop fashion in exactly the same way the single muscle system described at the beginning of this section. It is clear that since the target light which is to be pointed to is perceived visually, the central reciprocal command should be specified visually if the arm is to point at the target. In other words, the visual information about target direction must be translated into a reciprocal command⁹. The situation that has just been described illustrates a second way in which visual information can control activity: in an open-loop fashion setting the required goal state of the action at the beginning of the movement.

⁹ The shift in equilibrium posture of the limb in this task does not seem to be achieved by a step change in the central command, but by an evolution of the central command from an initial state to a final state corresponding to the required equilibrium posture (Bizzi et al., 1982, 1984). Thus the limb can be thought of as tracking a sequence of intermediate equilibrium positions as the central command evolves, the equilibrium trajectory hypothesis (Bizzi et al., 1984).

It is possible to augment the situation just described so as to fit in more of the known empirical data on the control of pointing movements in man. These will be briefly mentioned here for the sake of completeness. Firstly, the central (reciprocal) command can be made to depend continuously (or approximately continuously) on visual information about target position. This means that should the target move to another position during the execution of the movement the central command will change causing the limb to move to the new position of the target. This is in fact what happens (Soechting & Lacquaniti, 1983). Such modification of limb movement in response to changes in target position can take place without conscious intervention as indicated by the experiments of Prablanc, Goodale and their co-workers who changed the target position during the initial saccade to the target which lead to the human subject being unaware of the change but also to arm movement corrections (Goodale et al., 1986; Prablanc et al., 1986). Secondly, when the limb is visible pointing movements seem to be typically controlled in a closed-loop fashion by visual information about the relative position of the hand and the target though such control only comes into operation after about 200 milliseconds (for a review see Jeannerod, 1989). Both these extensions can be considered to be adding a visual feedback loop to the reciprocal commands which were described as open-loop above. The control scheme is now as illustrated in figure 2.5b.

The closed-loop system in this example can be considered to impose a dynamics on the state of the joint (as defined by the joint angle and its rate of change). These dynamics are spring-like and one may consider the behaviour of the limb under the control of this closed-loop system to be equivalent to its behaviour were it attached to the direction of the target light by a spring; the spring-like forces are due to articular proprioceptive information and visual information acts to tune the spring as a whole.



§2.6. PERCEPTUO-MOTOR CONTROL

In section 1 a control problem was defined as being specified by three items: (i) A system that is to be controlled (the plant). (ii) A specified pattern of behaviour that we require the plant to exhibit. (iii) Certain things (inputs, forcings or controls) that can be done to the plant which act to change its behaviour. The second item arises as a consequence of some specified task a person wants performed. The concern in the study of human and animal motor behaviour is with tasks which can be carried out in a variety of different ways by many physically distinct mechanisms. Therefore, it is useful to be able to describe such tasks in an abstract (mechanism independent) fashion. Since it is really a matter of commonsense that many tasks are independent of any physical mechanism which might carry them out, the notion that tasks should be characterised abstractly has been around implicitly for a long time. However, only recently has it appeared as an explicit part of any approach to biological motor control (see, for example, Arbib, 1975; Hollerbach, 1982; Saltzman, 1979; Saltzman & Kelso, 1987). The abstract descriptions of tasks adopted by these researchers were derived from work in robotics where it was found early on that such descriptions were extremely useful.

To make the discussion concrete a simple example will be considered, that of touching a particular point on a surface. Any device which performs such a task must move some physical structure such that it touches (without altering in any way) a specified point on a surface in its environment. The identity of the physical structure which touches the surface is unimportant, all that matters is that it is capable of playing the functional role of a "toucher" (a person could touch a surface with his finger, thumb, toe, nose, elbow, a stick or whatever). Similarly, the physical structure of the surface is of little concern; all that matters here is the position of the point to be touched in relation to the thing which is to touch it. Thus what is important about this task is simply the spatial relationship between two points: the environmental point to be touched and the point on the device which is going to touch it. Abstracting everything else about the situation away, deeming it irrelevant to the characterisation of the task, one is left with the elements of what is sometimes called a *task space* description in robotics (a term adopted by Saltzman & Kelso, 1987). In the 'touching' example this space simply represents the geometry of the relationship between that part of the functional device which actually performs the task (interacts with the environment), called the abstract

end-effector, and the point to be touched (figure 2.11). A performance of the task is represented simply a trajectory of the end-effector through the task space from its initial position to a final position at rest coincident with the point to be touched.

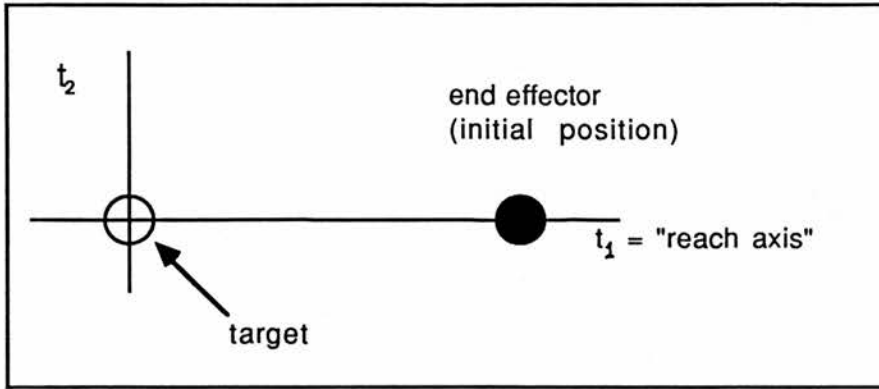


Figure 2.11. Task geometry of a touching task in two dimensions (after Saltzman & Kelso, 1987). The end-effector (that which touches the target) is modeled as an abstract point (filled circle). The target (open circle) defined the origin of a Cartesian coordinate system $t_1 t_2$.

The task space of figure 2.11 is the space of possible configurations of the end-effector, one of which represents the goal of the touching task. If a third dimension is added to the space to represent the velocity of the end-effector, then a space which represents the task relevant *state* of the end-effector/environment system has been constructed. One can view this state-space as that of a system which is to be controlled such that the specified goal state is reached. In control theoretic terms this is the state-space of a *plant*. Once a 'plant' has been defined for a task one is on familiar territory and straightforward control theoretic ideas can be applied. It follows, therefore, that a solution to the task can be represented as a dynamics defined on the plant's state space as usual.

The simplest dynamics to define on the state space of the touching task which represents a solution to the task is one of the form:

$$\begin{aligned} m\ddot{t}_1(t) + \beta_{t_1}\dot{t}_1(t) + k_{t_1}t_1(t) &= 0 \\ m\ddot{t}_2(t) + \beta_{t_2}\dot{t}_2(t) + k_{t_2}t_2(t) &= 0, \end{aligned} \quad (2.10)$$

where $[t_1(t), t_2(t)]$ is the position vector of the end-effector in the task space as a function of time; the β and k are, respectively, the damping and stiffness coefficients, and m is the effective mass of the end-effector in task space (cf Saltzman &

Kelso, 1987). The system has an equilibrium point at the origin of the coordinate system in which (t_1, t_2) is measured, which may be defined to be the goal position (the dynamics should have critical damping so that there is no overshoot of the goal position). This kind of dynamics is that of a linear spring in two dimensions and in the task space the end-effector behaves as if it were attached to the goal position by such a spring. Such a 2-D spring can be conveniently represented as two one-dimensional springs in a convenient coordinate system as in figure 2.12 below (Saltzman & Kelso, 1987).

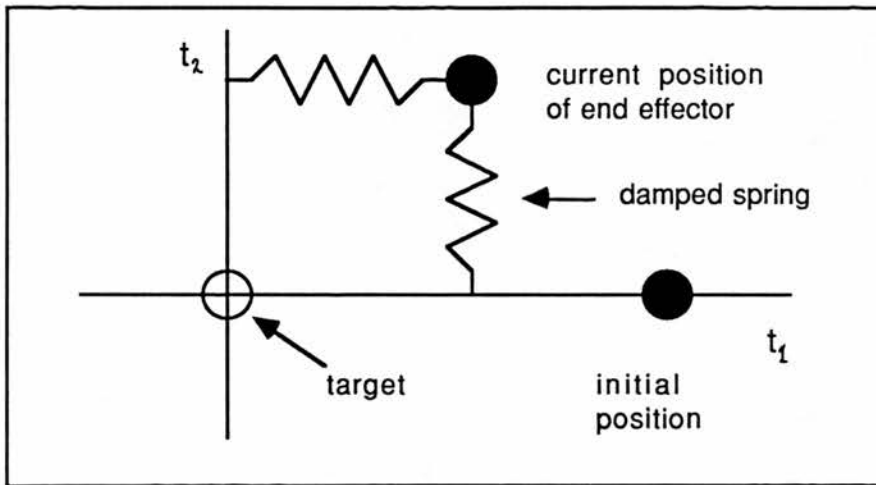


Figure 2.12. Task space dynamics represented as two virtual "springs" connecting the end-effector to the task axes (t_1, t_2) (after Saltzman & Kelso, 1987, figure 4).

The abstract representation of a task in terms of its task-space and its solution as a dynamics on a corresponding state space (which translates into a force field on the task-space, here represented by the action of springs) corresponds exactly to the analysis of the orientation behaviour of the housefly given by Reichardt & Poggio (section 2.4 above). The essential features of this kind of analysis can be stated as follows:

- A task is specified by a goal. Most goals such as switching on a light, cutting a slice of bread, touching a point on a surface and so on, can be achieved in many different ways (one could use similar movements but different tools or *vice versa* to achieve a given goal).

- Although a variety of different tools could be used in execution of a given task (different parts of an animal's body or different physical objects) they are all used for the same purpose — they are all *equivalent* in this respect. This functional equivalence is captured in the concept of an *abstract end-effector* which represents what is essential to the function of all tools which could be used to execute a task — the physical differences between various tools is abstracted away since it is not functionally relevant.
- To perform a given task a relationship between the end-effector and the environment must reach a desired state or state trajectory; this state or state trajectory is the goal of the task. The particular relationship that is relevant depends on the task, for example, in the case of the touching task considered above the relevant relationship was geometrical. For other tasks temporal and/or dynamic (forceful) relationships may be relevant.

By representing the goal as a single state in a space of possible states enables one to define an abstract “plant” (i.e., object to be controlled) whose states are precisely the states of this space. The plant in this case corresponds to a system defined by both the end-effector and the environment. The state of a relationship between the end-effector and the environment is a state of this plant.

- Finally, the control problem is to move the end-effector from any initial state to the goal state or state trajectory. A control law which solves this problem necessarily defines a dynamics on the state space of the abstractly defined plant. Thus, to understand how a task is performed by a biological system the task space dynamics need to be described. Such a description specifies the abstract control law that the biological system is implementing in performing the task. This is the message conveyed by the work of both Reichardt and Poggio (1976) and Saltzman and Kelso (1987) considered above.

The discussion presented so far in this chapter leads towards the conclusion that some form of analysis of the kind captured by the above scheme is a necessary first step if one is to understand how a system is being controlled so as to perform a task. This conclusion follows from the account of control developed above. As discussed, a controlled system performing a useful function is represented in scientific discourse as a dynamical system of some kind. Control theory is a discipline usually applied to the *design* of devices which achieve desired patterns of behaviour.

The desired behaviour is identified with a state trajectory of the plant (i.e., the

system of which we require the desired behaviour). The problem is then to *design* a control law which stabilizes the state trajectory representing the desired behaviour. Put in another way, the problem is to couple the plant to another system (the controller) such that the desired behaviour is an attractor (asymptotically stable state trajectory) of the combined system. In studying biological motor control one is faced with a control system whose operation is unknown. The problem is then not one of design but of *analysis*. One is thus working backwards and so must start by treating the goal of the task the system is executing as an attractor of its dynamics defined at level of functional behaviour and seek to give a characterization of this dynamics (simply a reversal of the strategy used in design). This provides a description of the action of the control system. Only when what control is achieving is understood can one hope to understand how it is implemented. This is analogous to the notion championed by Marr in vision research and closely related to the ideas of Gibson (see chapter 1), that the first step in understanding a perceptual system should be the precise understanding of *what* that system is doing.

§2.7. CONCLUSIONS

The conclusions that may be drawn from the above discussion are twofold. First, the dynamical systems approach and an approach based on computational and control theoretic concepts are not opposed but are complementary. As the example of flight control in the housefly illustrates, a dynamical analysis at the behavioural or task-space level is a necessary step in understanding the control mechanisms of the fly — what the fly is doing needs to be understood before one can hope to understand how it is doing it. The dynamical equations that describe the fly engaged in the act of orienting are a dynamical description of a (closed-loop) control system. Second, understanding the perceptuo-motor behaviour of animals requires a detailed understanding of the tasks that animals perform. A precise understanding of the goal of the task is required — the goal places certain constraints on an animal's motor behaviour, so-called task constraints (e.g., Newell, 1986). Understanding the task enables one to define a task-space and to determine the information about the state of the task space (the animal-environment system) that an animal will require if the task is to be performed effectively.

In the next chapter, simple interception tasks will be analyzed in an attempt to understand the information about the animal-environment system that is required

to perform such tasks effectively. In addition a model of a task-space control scheme for the timing of interceptions is presented which illustrates some of the concepts developed in this chapter. The remaining chapters seek to understand the nature of the information people actually use to time interceptions of moving objects. This is only half the story, of course, the other half, which is briefly addressed in chapter 3, is the dynamical characterisation of the human subject engaged in an interceptive action. This is beyond the scope of this thesis.

3

Control of Interceptive Timing

§3.1. INTRODUCTION

In the last chapter an attempt was made to provide a firm basis for adopting a style of analysis of perceptuo-motor behaviour in animals and man that has been advocated in recent years by researchers from the ecological school and to wed this approach to other current approaches. It was argued, following many others, that the tasks that an animal or person performs should be analyzed at an abstract level, independent of the precise mechanisms of execution, and that the information required for their successful performance be identified. The nature of the required information depends upon certain constraints, as will be illustrated below. In this chapter, tasks which involve interception of a moving target are considered (specifically catching and hitting) and the nature of the information required to effectively *time* their execution is investigated.

§3.2. INTERCEPTIVE ACTIONS

The performer of the action will be considered to be fixed as a whole relative to his environment (sitting or standing still) or moving as a whole relative to it. In these circumstances an *interceptive* action is one where the performer interacts with an object or surface which is in motion relative to him (the *target*): examples include making a shot in tennis, catching a ball or placing the feet in particular places whilst running. According to this definition, reaching out to pick up a stationary object whilst seated is not an interceptive action even though the object and the hand are in relative motion. An *avoiding* action is one in which the performer moves so as to avoid a collision with an object in motion relative to him: swerving and braking during running or driving are examples as is jumping over a obstacle whilst walking or running.

A moving object travels along a *trajectory*. This means that at any instant of time t the object is at some position in space. It is clear that in order to intercept such a moving object the intercepting end-effector (e.g., the hand) must, in some sense, be in the right place at the right time. In intercepting a moving target there are thus constraints both on where the end-effector must be and on when it must

get there if the task is to be carried out effectively. Since the target is in a given place at a given time the two constraints are not independent — when and where are mutually determining; if you select where an interception is to be made then the ‘when’ is determined. This much follows in a direct and straightforward fashion from the simple fact that the target is moving. An interception takes place at a point on a trajectory, at a spatiotemporal location. To specify a spatiotemporal location of the target in the environment four coordinates are required, three spatial and one temporal. How can information about a target’s spatiotemporal location at an instant of time or over an extended period be used to effect an interception and exactly what is the nature of the information that might be used? These questions are considered in the next section in the context of two interceptive actions, hits and catches.

§3.3. HITTING AND CATCHING

As an example, consider the task of hitting a moving target under the simple conditions illustrated in figure 3.1. The goal of the task may be expressed most simply as that of bringing the end-effector onto the path of the target coincident with the arrival of some part of the target (the end-effector must, of course, have a non-zero component of velocity perpendicular to the path of the target for the target to be actually struck). A hit is considered to have been made if some part of the target is contacted by some part of the end-effector. The ‘perfect’ hit may be considered to have been made when the point marked p on the end-effector contacts the centre of the target (marked c). It is clear that if a perfect hit is aimed for, small errors can be made and a hit still achieved since at the instant when the end-effector meets the path of the target some part of the target is in contact with some part of the end-effector.

To be able to perform such a hitting task the controller governing the motion of the end-effector needs information about the motion of the target so that it can tell the end-effector how to move so as to effect the hit. Exactly what is the nature of the information required? This depends in part upon the exact method that is used to execute the act. As described in chapter 2 it is almost always possible to execute any given action using a variety of different methods. The nature of the information required to effect the act can depend upon the method of execution which in turn will depend upon constraints on the possible motions of the end-effector and of the

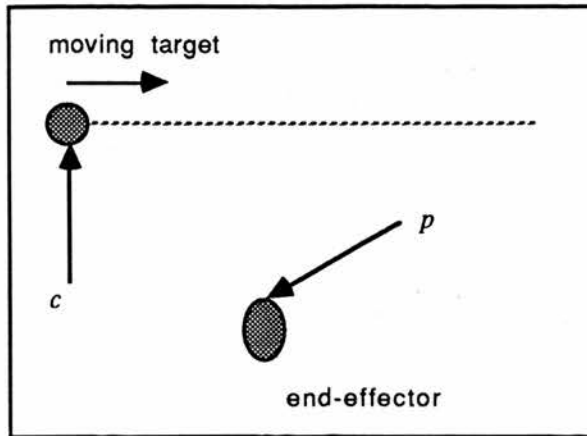


Figure 3.1. Simple geometry of interception.

target. Such constraints can arise from the following sources:

- (1) the environment in which the task is performed, e.g., objects and surfaces can prevent motion of the end-effector in certain directions.
- (2) The effector system, which may, for example, be a system of jointed linkages which can constrain motions of the end-effector to lie within a certain circumscribed region of space (the “workspace” as it is usually called in robotics).
- (3) The controller, which will have a finite computing capacity and a certain speed of response.

The hitting task described above will serve to illustrate how different methods for achieving the same end can influence the nature of the information required for its execution.

Suppose that the end-effector in figure 3.1 is some kind of missile. There are then no constraints on how far it can move from its initial (launch) position save running out of fuel. Consider two interception strategies that might be employed to hit the target: (1) a ballistic strategy (open-loop controller), (2) a guided strategy (closed-loop controller). The ballistic strategy requires that an interception point be determined and the missile fired accordingly. For the simple situation illustrated in figure 3.1 such a scheme might work as follows. Translate the situation illustrated in figure 3.1 into the idealized scene shown in figure 3.2. An object is moving along a straight horizontal path above the ground at constant speed. A missile launcher below on the ground is faced with the task of shooting it down. If the launcher

moves along the ground with the same horizontal velocity as the object, then there will be no motion of the object relative to the launcher. A missile need only be directed at the object and fired for the object to be shot down. If the launcher is fixed to the ground then the same result could be achieved by calculating the velocity of the object relative to the ground and firing a missile such that it has this velocity combined with another component directed at the object's position at the instant of firing. The missile is thus fired with a velocity which is the vector sum of these two components and it will hit the object at the point on its path marked c in the figure.

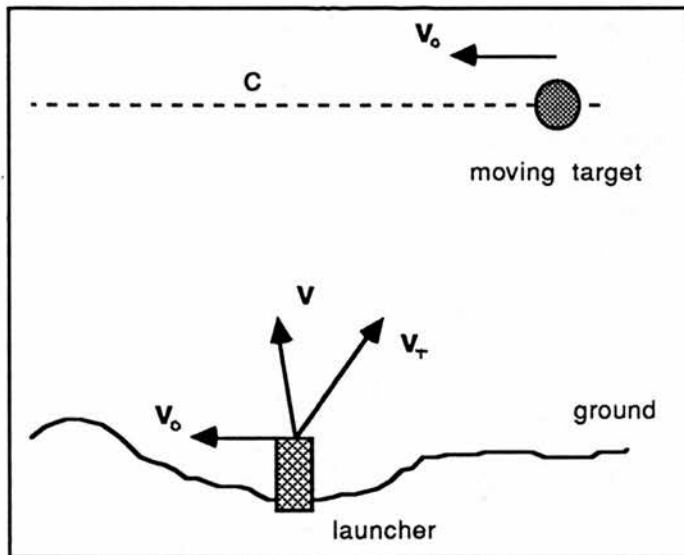


Figure 3.2. The "missile strategy". A missile launcher is on the ground and a target object is moving above it along a straight line path in the plane, shown as a dotted line. A missile fired with velocity v (the sum of a component v_o towards the instantaneous position of the moving target and a component v_o equal to the velocity of the moving target) will intercept the object at a point C along its path.

A version of the guided missile strategy could be implemented in a variety of ways. The simplest method would be to equip the missile with an on-board control system which measures the angular distance between the direction of travel of the missile and the position of the target and drives the missile so as to keep this angle equal to zero. Provided that the speed of the missile is greater than the speed of the target a hit will be achieved.

The two strategies are both capable of achieving interception of a moving target whose speed is not too great, but the information about the target they require is different. The ballistic strategy requires information about both the *direction* of the target and its *velocity* relative to the launcher. The simple guided strategy requires only information about the target's direction relative to the direction of missile flight.

Strategies similar in principle to the above have been proposed as accounts of how human performers might catch moving targets. A catch implies additional task constraints to those that must be satisfied by a hit: the moving target must be gripped and held by the end-effector which will typically also be constrained to remain within a circumscribed workspace thus imposing the requirement that the velocity of the target relative to the frame of reference of the workspace be reduced to zero before the end-effector leaves the workspace. The ballistic strategy is the basis of the account proposed by von Hofsten (1983, 1987) for the ability of human subjects¹ to catch target objects moving along simple paths such as straight lines. Von Hofsten takes the information about the target to be provided by vision and proposes that the motion of the hand to the target is based on the above ballistic strategy. Thus, when vision is only available to control the hand's motion prior to the initiation of the action, von Hofsten's strategy is identical to the ballistic strategy. Under normal circumstances, when visual information about the target is available throughout the act, von Hofsten proposes that this information is used in a feedback control fashion to guide the hand (end-effector) to the target (he does not, however, provide a control law for implementing this idea so it is not clear exactly how the proposed control is supposed to work). Notice that the ballistic strategy only works for targets moving along straight line paths. If this kind of approach is to be successful for targets moving along curved paths a feedback modulation is required.

In a little known paper, Chapman (1968) describes a feedback control scheme which he proposes as a method by which an outfielder might catch a baseball. The baseball moves on a projected free-fall trajectory under gravity and neglecting air resistance and other aerodynamic effects, the path of the ball is a parabola and motion is described by Newton's kinematic equations. Assuming the coordinates

¹ In the study reported in von Hofsten (1983) the subjects were young infants.

in figure 3.3 one can write

$$\begin{aligned}y(t) &= Vt \sin \theta - 1/2gt^2 \\x(t) &= Vt \cos \theta.\end{aligned}\tag{3.1}$$

The range R of the ball is given by

$$R = \frac{2V^2 \sin \theta \cos \theta}{g}.\tag{3.2}$$

Suppose a person wishing to catch the baseball (a fielder) is standing a distance R from the point of the ball's projection (i.e., at point F in figure 3.3). At an instant of time t the ball will be at an elevation angle $\phi(t)$ as seen by the fielder. From the geometry in figure 3.3 we have that

$$\tan \phi(t) = \frac{y(t)}{R - x(t)},\tag{3.3}$$

substituting for x and y from equation 3.1 and for R from equation 3.2 one finds that

$$\tan \phi(t) = \left(\frac{g}{2V \cos \theta} \right) t.\tag{3.4}$$

Under the assumption that aerodynamic effects can be neglected, the term in the brackets on the right hand side of equation 3.4 is simply a constant. Thus, for a fielder at the point F , the tangent of the ball's elevation angle increases at a constant rate over time until it reaches him. Chapman (1968) shows that if the fielder is standing anywhere else the tangent of the angle of elevation does not increase at a constant rate. The rate of increase of $\tan \phi$ for three possible fielder positions is shown in figure 3.4.

On the basis of the above analysis Chapman suggested a simple feedback control law which will get the fielder to the right position at the right time (provided he is standing initially in the ball's plane of motion²). The idea is for the fielder to move such that the rate of increase of $\tan \phi$ is a constant. How can the fielder do this? Chapman suggested the following strategy. He assumes that the fielder is at some initial position a distance s from F in figure 3.3 and runs at a constant speed in the right direction (forward or back). He knows which direction to run in because if $\tan \phi$ changes according to the function below the line of constant increase he must run towards the ball and if it lies above this line he must run away from the

² If he is not standing in the right plane to start with one could apply a second control law which tended to bring him onto the correct plane as noted by Chapman.

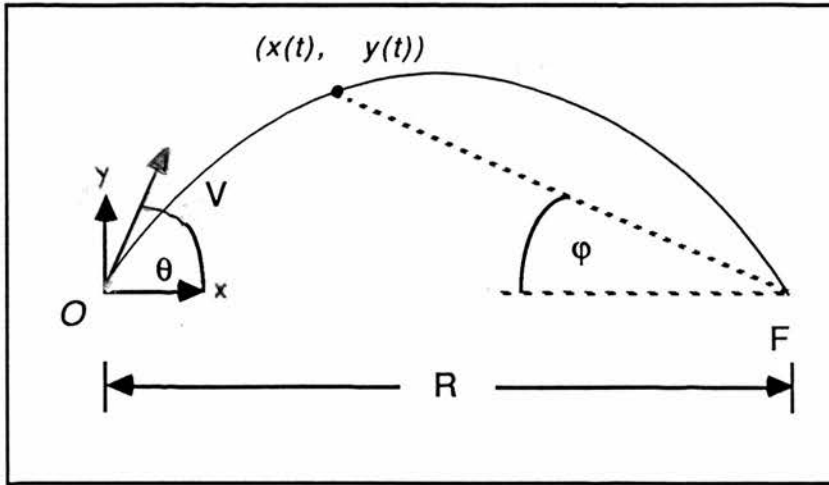


Figure 3.3. A ball is projected from O (hit by a batter) with velocity V and travels along the curve shown returning to its original height at a point F a distance R from O . At some instant of time t the ball is at a point $[x(t), y(t)]$.

ball. Running at constant speed towards or away from the ball is equivalent to increasing or decreasing the horizontal velocity of the ball relative to the observer. The observer's task is to find a speed which makes the rate of increase of $\tan \phi$ constant. This will be difficult to achieve since the time taken to assess whether or not the rate of increase of $\tan \phi$ lies above or below the constant line will increase as $\tan \phi$'s rate of increase approaches a constant. The reason for this is evident in figure 3.4: the curves do not significantly diverge until some period of time has elapsed and this period depends on how close the observer is from the point F . It is this divergence that the observer must detect and decide which way the current curve is diverging from the line of constant increase.

Neither von Hofsten's nor Chapman's proposals can account for the catching skills of human performers for two reasons. First, they both ignore the definitive aspect of the act of catching since they fail to address the problem of how the target object is gripped by the catcher. It is not enough simply to get to the right place at the right time if one wants to effect a catch rather than a hit. This can be most easily seen by considering a two handed catch made between the palms of the hands as often attempted by young children (Williams, 1985). The hands start some distance apart and must close on the target to effect the catch. If the hands close too soon by the time the target reaches them the aperture will be too narrow for the target to pass through and the catch will be unsuccessful. If the

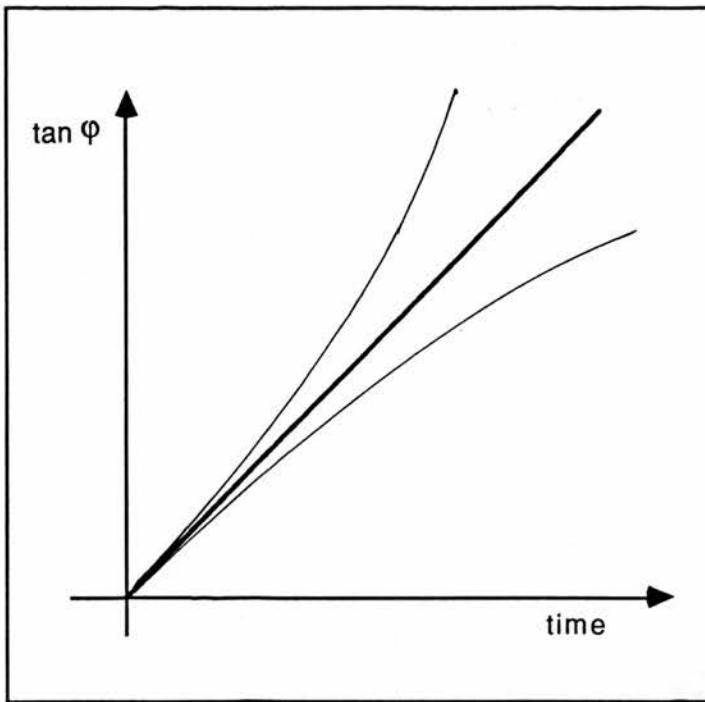


Figure 3.4. The rate of increase of $\tan\phi$ for three possible fielder positions. The lower curve is for a fielder between F and the projection point, the straight line is for a fielder at point F and the upper curve is for a point beyond F in figure 3.3.

hands close too late the ball will pass through the aperture before it can be gripped. In order to perform such a gripping action so as to effect the catch, information about the time when the target will pass through the aperture is required (a time-to-contact). The controller is not provided with such information in either von Hofsten's or Chapman's catching schemes, indeed, von Hofsten explicitly states that his scheme obviates the need for explicit timing information, an assertion which cannot be justified.

The second reason for the failure of the proposals to account for human catching skills is that they assume that the intercepting end-effector is not confined to lie within a circumscribed workspace. Consequently they cannot account for the familiar one-handed catches often executed by people in which the catcher simply reaches out and catches, for example, a ball which is thrown to him or her. Chapman's scheme fails in this case because the observer's eye must be able to move unrestrictedly. The same is not true for von Hofsten's scheme which fails for another reason. Assuming that the performer is positioned such that the target passes through his workspace the catch must be effected whilst the target is within this space, by definition. The performer is constrained to move the end-effector from its initial position and velocity (which will often be zero) in the work space

to some interception point. In addition the end-effector's velocity in the direction of the interception point must be brought to zero before the boundary of the work space is reached (typically this velocity will be zero at the moment of the catch itself). These constraints mean that the performer cannot employ von Hosten's simple missile scheme to catch the target because this scheme assumes that the end-effector velocity is more or less constant. Referring back to figure 3.2, to calculate the intercept velocity v_m one needs to have a constant missile velocity v_T . One could imagine a scheme based on the average velocity of the end-effector, but any average depends upon the distance to be moved and the time spent in motion since

$$\text{average speed} = \frac{\text{distance to be moved}}{\text{time spent in motion}}$$

The time spent in motion is none other than the time-to-contact of the target with the interception point. Thus one could compute an appropriate average speed using information about the distance to be moved and the time-to-contact. It appears that for a performer constrained in the manner described, time-to-contact information is required for successful catching. The same is true for hitting if this is construed as the task of making an end-effector contact a moving target when the end-effector is constrained to stay within a circumscribed work space and an optimal hit is achieved when the contact occurs at the moment the end-effector attains its maximum velocity. This task is discussed in more detail in section 3.5 below.

Conclusions. Interception of moving targets requires that the control system which drives some end-effector to make contact with the target has information about the trajectory of the target. In the examples described here the trajectory is determined by the target's position and velocity at any instant of time, and it is possible to achieve interception based on information about these two quantities alone as the missile strategy of von Hofsten illustrates. However, when the interception is constrained to be effected within a circumscribed workspace by an end-effector which accelerates from its initial state and must stop before it reaches the work space boundary one needs to use a third quantity; the time-to-contact of the target with the interception point. This is determined by the position and velocity of the target at an instant and the position of the interception point and might be computed from these quantities. In later chapters the question is addressed as to

whether the processes involved in human control systems actually perform such a computation or obtain time-to-contact information in other ways.

We have seen here how time-to-contact information is important in catching an hitting. It is possible that time-to-contact is also important in a variety of other interceptive actions and Lee has argued for the use of time-to-contact information in the control of activities such as placing the feet during locomotion over rough ground (Lee, 1980; Warren, Young & Lee, 1986), assuming the right posture for entry into the water when diving (Lee, 1980; Lee & Reddish, 1981) and running and jumping (Lee, 1980; Lee, Lishman & Thompson, 1982; see also Warren & Kelso, 1985).

§3.4. THE CONSTANT VELOCITY STRATEGY IN INTERCEPTIVE TIMING

Lee (1980) introduced the idea that even in the presence of relative acceleration between an animal and an object (or surface), the animal might time an interaction with such an object on the assumption that the velocity is constant. Lee dubbed this strategy the tau-strategy (Lee & Reddish, 1981; Lee et al., 1983) since this is the strategy an animal would follow were it using an optical source of time-to-contact information called tau (defined in the next chapter) to time interactions with moving objects. In order to avoid confusion, it should be stressed that Lee's tau-strategy is independent of whether or not this optical source of time-to-contact is actually used in interceptive timing; for this reason it might better be called simply the *constant velocity strategy* and it will be referred to by this name below.

A small but significant body of empirical results support the notion that a constant velocity strategy is adopted in interceptive timing. First, Lee has shown that gannets, plunge diving to catch fish, follow such a strategy in timing the retraction of their wings before entering the water (Lee & Reddish, 1981). In another study he presents evidence that human subjects leaping to punch a ball falling towards them (thus accelerating under gravity) follow the constant velocity strategy (Lee et al., 1983). Second, Todd (1981) measured the threshold for visually detecting differences in time-to-contact in computer generated images and found that observers were unable to take accelerations into account effectively. Third, several studies indicate that human observers are insensitive to object acceleration perpendicular to the line of sight (Gottsdanker, 1956; Gottsdanker, Frick & Lockerd, 1961; Runeson,

1975). Gottsdanker et al. (1961), for example, found that smoothly accelerated motion was not well discriminated from unaccelerated motion. This effect was most marked for brief viewing times (less than one second) — with longer exposure discrimination improved. Gottsdanker et al. interpreted these results to mean that acceleration is not identified by “direct sensing”, which may be taken to mean that information specifying target acceleration is apparently not being used by the subjects in their judgements. Subjects’ judgements were consistent with a strategy of discriminating accelerative motion on the basis of the target having different velocities at different times. (Noticing that a target has a different velocity at one time than at another is not the same as perceiving acceleration.) These latter results provide indirect corroboration of a constant velocity strategy in interceptive timing by showing that where information about object acceleration might be used it apparently is not.

Lee has also pointed out that the constant velocity strategy can be used effectively to time interceptions during accelerative approach (Lee et al., 1983) in a simple and robust fashion making it unnecessary to extract more complicated information which takes acceleration into account (this point is further discussed in Appendix A). Although the cited evidence is consistent with a constant velocity strategy, it is important to note that *direct* evidence for such a strategy in interceptive timing has only been provided for the case of collision (or approximate collision) with the observer’s eyes (Lee & Reddish, 1981; Lee et al., 1983). No direct evidence for the constant velocity strategy for non-collision trajectories has been published.

§3.6. CONCLUSIONS

In this chapter we have examined abstract simplified versions of basic interceptive tasks (catching and hitting) which involve controlling collisions with moving target objects. The basic constraints on these tasks were analyzed and it was concluded that effective timing requires information about the time-to-contact of the target object with a proposed interception point contrary to the opinion of von Hofsten who has argued that time-to-contact is not required for catching and hitting tasks. The work of Lee and others suggests that in timing interceptive actions, humans as well as other animals do not employ the actual time-to-contact but employ an approximation to it based on the assumption that the target is moving with a constant velocity towards the interception point. This is the constant velocity strategy in interceptive timing also called the tau strategy (Lee et al., 1983).

The analysis presented in this chapter shows how investigation at the level of description of the task can serve to identify the information required to execute a task and formulate models of control based on this information. In the next two chapters the question of how the required information might be obtained by a person or animal is investigated. In chapter 4 the question of what information about time-to-contact is visually available to an observer is examined. In chapter 5 questions concerning how such information might be extracted from the visual stimulus are dealt with. The constant velocity strategy is assumed throughout (but see Appendix A).

4

Visual Stimulation and Information

§4.1. INTRODUCTION

In the present chapter we address the problem of describing the information present in visual stimulation which subserves the visual perception of an environment and visual guidance of action. The particular concern is with information present in the visual stimulus of a monocular observer due to *motion*: either motion of an observer in a *rigid* environment or motion of objects through the environment relative to the observer.

In order to specify precisely the problems involved in the description of visual information and in its extraction by visual systems it is necessary to define and distinguish certain concepts that appear in the literature analyzing these problems. We begin, therefore, by discussing the nature of visual stimulation. The concepts of optic flow, image flow, optic velocity field and image velocity field, often used in the literature interchangeably, are distinguished in order to highlight certain problems of information extraction or “pick-up”. In consequence, the definitions given differ slightly from some that have appeared elsewhere. The problems of information extraction are discussed in more detail in chapter 5. The question of what information is present in the visual stimulus is addressed largely through an analysis of the (monocular) *image velocity field* which serves once again to emphasize problems of information extraction. An overview of the information present in the image velocity field is presented together with a detailed analysis of available information for timing activities involving interaction with an environment in motion relative to the actor. In this context the “optic” variable tau (τ , actually a velocity field variable) introduced by Lee is defined. Certain restrictions are necessary for tau to be useful in defining timing information. The rigidity constraint has already been mentioned. In addition, for target objects moving through a rigid environment use of tau in timing also requires that either (i) the object is on a collision course with the observer and not rotating or (ii) the surface of the object is normal to the line of sight and not rotating or (iii) the object is a sphere. In this chapter it will be assumed that the relevant conditions are met. Chapter 5 will examine the consequences of relaxing some of these conditions.

§4.2. AVAILABLE VISUAL STIMULATION

The visual systems of animals and the visual mechanisms of machines are stimulated by light falling on receptive surfaces. The human eye, like the eyes of most animals, focuses light convergent at a point but not occluded by the casing of the eye, as an image on a retina. No retinal image as such is identifiable, however, in the compound eyes of arthropods. The fact that an identifiable retinal image is not an *a priori* necessity for vision and that only light not occluded by the imaging device is focused as an image prompted Gibson (1950, 1966) to introduce the concepts of the *optic array* and the *optic flow*. These concepts describe the available visual stimulation at a stationary and a moving point of observation respectively, in a fashion that is *independent of any particular visual system that might actually be stimulated*. These concepts will now be elucidated according to Gibson's usage.

The Optic Array. An (ambient) optic array is defined at any point in an illuminated environment that can serve as a point of observation, i.e., a point where an act of observation could be made or an observer might be—essentially any point within a transparent medium such as air or water. (As such the optic array is 'monocular'). An arbitrary point of observation will be denoted \mathcal{O} and the sheaf of light rays that converge on \mathcal{O} is the optic array. Typically, rays will converge on \mathcal{O} from all directions. The set of all directions from \mathcal{O} (called the manifold of visual directions by Koenderink, e.g., 1985) clearly subtends the same solid angle (4π steradians) as a sphere centered on \mathcal{O} . Every point on this sphere is thus in a unique (visual) direction and the points on the sphere can be used to label the visual directions—the visual directions can be parameterised by spherical coordinates. For this reason it is convenient to conceive of the optic array as a pattern of light playing over the surface of a sphere centered on the observation point.

As emphasised by Gibson such an array is densely structured or patterned, indeed, if the light ambient to a point had no structure it would fail to qualify as an array at all—it has no *arrangement* (Gibson, 1979). The structure of an array is a consequence of the fact that the light is reflected from the (textured) surfaces that bound the solid and liquid materials making up the environment. It is in virtue of this structure that the array carries information about the environment. The environment can be considered as a complex layout of surfaces of varying extent, overlying each other and nested within each other in a complex hierarchical

fashion. The environment has identifiable geometric and textural structure at many different scales of spatial resolution. There will also typically be events occurring at different spatial scales and over different time scales. Gibson (1950, 1966, 1979) and Marr (1982) both describe in some detail the structure of physical surfaces and surface layout that gives rise to structure in visible light. Roughly, every unoccluded surface will subtend a solid angle at \mathcal{O} which will be filled with rays reflected from that surface. Most surfaces will be densely covered in *texture elements*—surface patches which reflect light differently from neighbouring surface patches due to different reflectance properties, different orientation or both. Such elements will thus be optically differentiable from their neighbours and in the optic array the borders of texture elements, objects and surfaces will tend to be marked by more or less abrupt changes in light intensity. The (instantaneous) structure of the optic array will take the form of a nested hierarchy of bounded visual solid angles (Gibson, 1979)—the ‘images’ of objects, surfaces and texture elements.

Differing reflectance properties and illumination conditions of different surfaces, surface regions and texture elements guarantee (given the physics of the propagation of light) that the important changes in the scene (the boundaries of spatially separate surfaces, changes in orientation of surfaces etc.) will be marked by changes in the optic array. The array is a projection of the scene but, as Marr (1982) observed, light intensity changes can be due to one or more of the following factors: the geometry of the scene, surface reflectances and illumination conditions. There is a significant problem in deciding which changes are due to which factors or to what combination of factors. A projection of the scene geometry can thus be said to be *implicit* in an optic array.

Optic Flow.* Suppose now that the observation point \mathcal{O} is moving (through open space) relative to an illuminated environment. It will travel along a path, at each point of which there will be a unique optic array. As it moves, therefore, it ‘occupies’ a different optic array at each instant and the light converging on \mathcal{O} will be changing continuously over time. The continuously changing sheaf of rays incident at a point of observation induced by a motion of this point through ambient space is Gibson’s optic flow. If the manifold of visual directions is conceived of as a spherical surface centered on \mathcal{O} , then the optic flow is the flow of light over this surface.

It is worth noting that in the *natural* world optic flow defined in this way only occurs due to translation of \mathcal{O} through the environment. This is because the frame

*The definition of optic flow given here and used throughout this chapter is the original one due to Gibson. It is important to note, however, that modern usage of the term, notably within the computer vision community (see, e.g., Horn, 1986), corresponds to what is here called *image flow*.

of reference of the whole visible scene is always the same as the inertial (mechanical) frame of reference of the environment. These two frames can be decoupled by artificial means but under such circumstances an animal or person tends to react as if the two frames were still coupled. For example, when the whole visible scene is rotated about a person (e.g., in an optokinetic drum) an experience of being mechanically rotated within a stationary environment is rapidly induced, so-called "circularvection" (Mach, 1875; Fischer & Kornmüller, 1930; Howard, 1982). In a similar way, when the visual scene translates as a whole relative to the observer an experience of translatory motion is induced (Fischer & Kornmüller, 1930; Wood, 1895) and compensatory postural movements are observed (Lee & Aronson, 1974; Lishman & Lee, 1973; Wood, 1895).

Optic flow as defined above is restricted to changes in the light incident at a point of observation induced by its motion through a rigid environment. Changes in the optic array due to local changes in the environment as occur, for example, when animals move through the environment or the leaves and branches of trees or other plants are moving in the wind, might perhaps be called *optic* or *visual motions* to avoid confusion. The term 'optic flow' will be reserved for optical changes due to motion of the observation point. This is in accord with common usage (Gibson, 1966, 1979; Marr, 1982). According to this usage it is possible to have optic motion within the optic flow, e.g, when both the observation point and surfaces in the scene are moving relative to the inertial environment. There is clearly a problem involved in determining the optical changes which are part of the flow (selfmotion) and those which are due to object motion.

Images and Image Flows. If the centre of projection of an imaging device, such as a camera or an eye, is placed at \mathcal{O} , then the light rays converging on \mathcal{O} which are not occluded by the imaging device itself are focused as an image on a surface. The observation point is the centre of projection. If the device is an animal's eye, the imaging surface is the retina. A retinal image is thus a projection of some portion of the optic array onto the retinal surface of an eye.

When an eye rotates about a point of observation¹ the portion of the optic array imaged on the retina changes; Gibson spoke of the retina moving 'under' the optic array and sampling different regions of it. Clearly the optic array does not

¹ The centre of rotation of the human eye doesn't quite coincide with this point but the difference is small and will be ignored here.

change as the eye rotates—it is a projection of the scene and the scene itself is independent of any rotation of an eye or imaging device. However, eye rotation leads to a continuously changing retinal image—an *image flow*. Image flow is thus optical motion relative to the imaging surface. Positions on such a surface can be labelled using a coordinate frame fixed to it—a retinocentric frame. As the imaging surface rotates about the centre of projection the coordinates of points in the imaged optic array measured in this frame clearly change. In general, therefore, *image flow is motion of the optic projection relative to a coordinate system fixed to the imaging surface (retina)*.

When the eye moves through the environment a portion of the optic flow generated by this motion will be imaged on the retinal surface. There will thus be a retinal flow due to the translation of the eye through the environment and if the eye is rotating there will be a component of retinal flow due to this rotation. It is worth noting that eye rotation relative to the environment can be due to either eye rotation about the optical centre (rotation of the eye in the head) or due to rotation about an axis external to the eye (the eye remains fixed relative to the head).

The rotational components of the retinal flow are not contained in the optic flow as defined above. The optic flow corresponds to the component of retinal flow due to translation, a retinal flow can thus have two components: an optic flow component and a component due to eye rotation (this will be discussed further in §3 below). The eye images the optic flow only if the retinal coordinate frame does not rotate relative to the environment, i.e., the orientation of the retinal coordinate frame must remain constant relative to a frame of reference fixed to the environment. This serves to illustrate an important fact about the optic array and the optic flow in the sense that Gibson used:— although the optic array (and flow) do not have any intrinsic coordinates associated with them, any coordinates that may be introduced in order to give a description of the optic array or flow cannot rotate relative to the frame of reference of the environment (Gibson considered the optic array at a point of observation to be anchored to the environmental layout). This fact is overlooked by Cutting (1986) who proposes that optic flow coordinates can rotate relative to the environment (see his figure 11.4 pp. 173).

Velocity Fields and Visual Flows. Almost all analyses of the information contained in visual (optic or image) flows analyze an *instantaneous* representation of the flow

called the *velocity field* (e.g., Clocksin, 1980; Gibson, Olum & Rosenblatt, 1955; Lee, 1974; Longuet-Higgins & Prazdny, 1980; Koenderink, 1986; Koenderink & van Doorn, 1975, 1981, 1987; Prazdny, 1983; Subbarao, 1988; Waxman & Ullman, 1988). Such velocity fields are *abstractions* from the available spatio-temporal variations in light intensity that constitute the actual visual stimulation (the visual flows themselves), a statement that requires explanation.

The velocity field representations that have appeared in the literature are two-dimensional projections of the instantaneous velocities of points on visible surfaces (relative to a coordinate frame fixed to the surface of projection). Effectively the velocity field is the 'image' of the motion of geometric points on the visible surfaces. It is thus necessary to distinguish a velocity field representation of a flow from the flow itself. The latter is a time-varying light intensity pattern that clearly depends upon the illumination of the scene, reflectances of surfaces and the physics of light; the former, in contrast, is a field of velocities which does not depend upon illumination, reflectance or physics. The velocity field depends only upon the time-varying geometry of the scene.

The difference between the velocity field and the instantaneous visual flow serves to define one of the problems of information 'pick-up'. If the information that has been found to be present in velocity field representations of instantaneous visual flow is to be picked-up by a visual system it must be possible to obtain the velocity field from the flow. This is a step that has tended to be overlooked in ecological psychology (Gibson, 1979). In fact, even the assumption that the velocity fields that have been used to represent visual flow are available at all in the directly given temporally varying light intensity patterns is a significant one. The assumption is valid if spatial variations in light intensity correspond to the texture elements, features and boundaries of visible 3-D surfaces (cf. Hildreth, 1984; Marr, 1982). In the discussion presented here this assumption will be made and as discussed above there is good reason to suppose that it will be true to a large extent in everyday environments (Gibson, 1979; Marr, 1982). It should be borne in mind, however, that such an assumption is known not to be unrestrictedly valid (Horn, 1986; Verri & Poggio, 1986). The problem of obtaining a velocity field from a flowing pattern of light intensities will be returned to briefly in chapter 5 where the problems of information pick-up are discussed.

It is worth noting that the visual flows that constitute the stimuli for real visual

systems will be extended over considerable periods of time and are not simply instantaneously defined entities like the velocity field representations that are widely analyzed in the literature. As a consequence of this there may be more information in the flow than is carried by the velocity field. Gibson (1966, 1979), for example, made the suggestion that certain information may only be defined over extended periods of time—it seems clear, for example, that since certain spatio-temporal happenings or “events” take an extended period of time to occur and hence any visual information about them will evolve over this period (Johansson, 1950; Johansson, von Hofsten & Jansson, 1980). Another example is the information about the direction of egomotion. As discussed by Warren, Morris & Kalish (1988) the direction of translational heading during curvilinear movement of an observer through a rigid environment is not present in the instantaneous velocity field which contains at most the instantaneous direction of movement which is clearly not constant over time during curvilinear motion. Another way of putting this is to say that the two rotational motions (rotation about an axis through the observation point and rotation about an external axis) are *instantaneously* indistinguishable². To obtain the direction of heading on a curvilinear trajectory the flow needs to evolve over some extended period of time or higher order accelerative components of the flow are required (Rieger, 1983; Warren et al., 1988). A third example is that of the perception of 3-D “structure from motion” (cf. Ullman, 1979). Perceptual studies indicate that the human visual system requires an extended period of time to accurately perceive the 3-D structure of a moving object when only motion information is available (Wallach & O’Connell, 1953; White & Mueser, 1960; Doner, Lappin & Perfetto, 1984; Ullman, 1984).

² This follows from a mathematical result, much employed in studies of the mechanics of rigid bodies, known as Chasles’ Theorem (Whittaker, 1944). This theorem states that any motion of a rigid body can be instantaneously represented as the sum of a pure translation and a rotation about an axis through the centre of the body.

§4.3. INFORMATION IN THE VELOCITY FIELD

The concern of this section is the problem of describing the information present in visual flows represented as instantaneous velocity fields. The discussion will be restricted to flows generated by motion of the observer relative to a *stationary* (rigid) environment—the case that has been most widely studied in the literature and that is of most use in the analysis of timing information. This is part of the problem known as the *interpretation of visual motion*³ (cf., Longuet-Higgins & Prazdny, 1980; Marr, 1982; Subbarao, 1988; Ullman, 1979). It is essentially the problem of showing how the time-varying geometric structure of the velocity field generated by motion of an observer through an environment can specify the geometry of that environment and how the observer is moving relative to it. As Waxman & Wohn (1988) have observed, the interpretation problem is a study in time-varying projective geometry.

The Image Velocity Field. In order to describe an image velocity field an imaging surface coordinate system must be introduced. Most commonly a planar model imaging surface is adopted (e.g., Longuet-Higgins & Prazdny, 1980; Subbarao, 1988). This does not compromise the generality of the analysis—there is a one-to-one correspondence between an image or image flow on a curved surface and one on a plane surface (cf., Lee, 1974; Longuet-Higgins & Prazdny, 1980), the shape of the imaging surface is simply a matter of mathematical convenience. The eye can thus be modeled by the situation illustrated in figure 4.1—Longuet-Higgins & Prazdny's "camera" model, a planar "retina" and associated coordinates.

The following derivation of the image motion of a point in the scene due to motion through a rigid environment follows, to a large extent, that of Longuet-Higgins & Prazdny (1980). In the camera model illustrated the observation point \mathcal{O} forms the origin of a cartesian coordinate system \mathcal{OXYZ} . The Z axis passes through the centre of the image plane \mathcal{o} which lies at $(0,0,1)$ in \mathcal{OXYZ} coordinates. The point \mathcal{o} forms the origin of an image coordinate system \mathcal{oxy} which is aligned such that the x and y axes are parallel to the X and Y axes respectively. Arbitrary rigid body motion of XYZ relative to the scene can be instantaneously represented as a translation \mathbf{V} and a rotation \mathbf{R} about \mathcal{O} (see footnote 2). These vectors are

³ Ecological psychologists object to the use of the term "interpretation", however, as was discussed in chapter 1 there is really no substance to this objection and use of the term is simply a matter of taste.

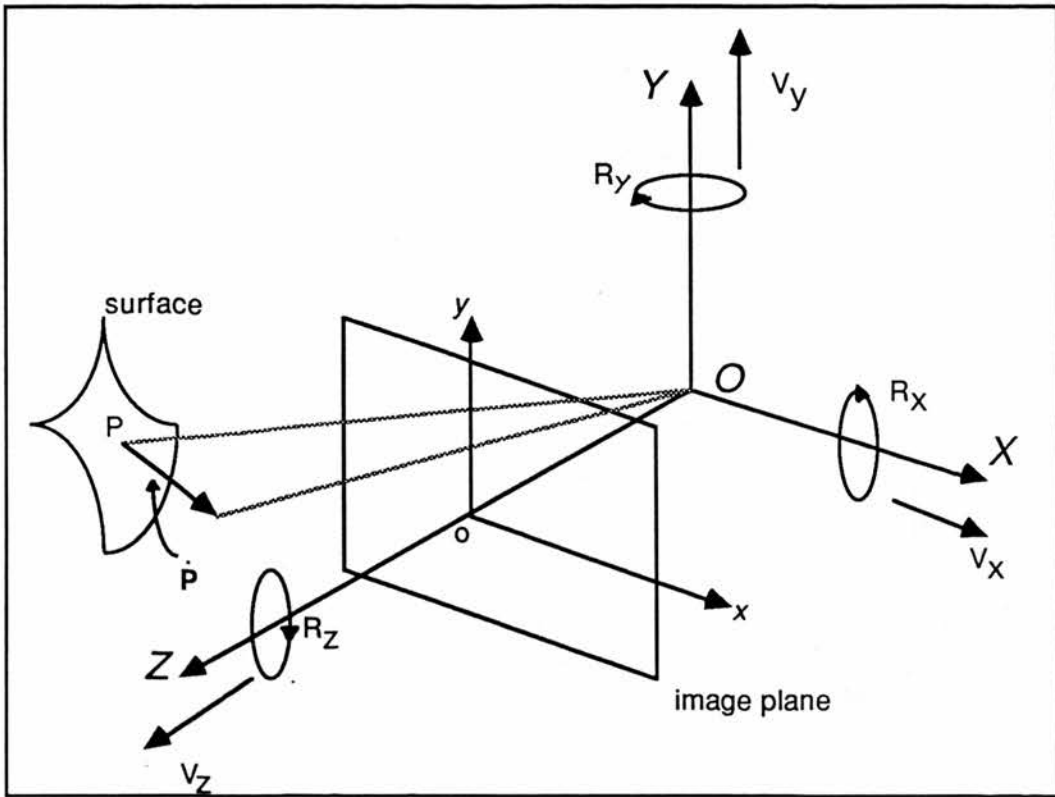


Figure 4.1. 'Camera' model imaging system, after Longuet-Higgins and Prazdny (1980). O is the centre of projection which forms the origin of the Cartesian coordinate system $OXYZ$, with Z being the direction of view. The image plane is fixed a unit distance along the Z axis in front of the centre of projection and oriented parallel to the XY plane: oxy is the coordinate system of the image plane.

represented in component form in figure 4.1: \mathbf{V} has X , Y and Z components V_x , V_y and V_z respectively, \mathbf{R} has components R_x , R_y and R_z ⁴.

If the instantaneous position vector of the point \mathcal{P} relative to $OXYZ$ is \mathbf{P} having coordinates (X, Y, Z) , then its instantaneous velocity vector is $\dot{\mathbf{P}}$ or, in coordinate form $(\dot{X}, \dot{Y}, \dot{Z})$. Clearly $\dot{\mathbf{P}}$ is determined by motion of the camera as described by \mathbf{V} and \mathbf{R} , in fact we can write⁵

$$\dot{\mathbf{P}} = -(\mathbf{V} + \mathbf{R} \times \mathbf{P}), \quad (4.1)$$

where \times denotes the vector product. (4.1) is usually written out in component

⁴ All these quantities may vary in time and should thus be strictly written as functions of time. However, since the instantaneous case is being considered we shall not explicitly write the quantities as functions of time as this can be understood.

⁵ This is straightforward application of the Coriolis Theorem, see, for example, Woodhouse (1987).

form:

$$\dot{X} = -V_X - R_Y Z + R_Z Y \quad (4.2a)$$

$$\dot{Y} = -V_Y - R_Z X + R_X Z \quad (4.2b)$$

$$\dot{Z} = -V_Z - R_X Y + R_Y X. \quad (4.2c)$$

The projection of the point \mathcal{P} is denoted p in fig. 4.1 and has (instantaneous) image plane coordinates (x, y) which are given by

$$x = X/Z \quad \text{and} \quad y = Y/Z. \quad (4.3)$$

Alternatively, one could state that the image at point (x, y) on the image plane corresponds to the point (xZ, yZ, Z) in the scene. The instantaneous velocity of the image point $p(\dot{x}, \dot{y})$ is the projection of $\dot{\mathbf{P}}$ onto the image plane and can be obtained by differentiating equations (4.3), which gives

$$\dot{x} = \frac{\dot{X}}{Z} - \frac{X\dot{Z}}{Z^2} \quad \text{and} \quad \dot{y} = \frac{\dot{Y}}{Z} - \frac{Y\dot{Z}}{Z^2}. \quad (4.4)$$

Substituting for \dot{X}, \dot{Y} and \dot{Z} in (4.4) from (4.2a-c) the following expressions for the image velocity components are obtained

$$\begin{aligned} \dot{x} &= \left(\frac{xV_Z - V_X}{Z} \right) + [xyR_X - (1+x^2)R_Y + yR_Z] \\ \dot{y} &= \left(\frac{yV_Z - V_Y}{Z} \right) + [(1+y^2)R_X - xyR_Y - xR_Z]. \end{aligned} \quad (4.5)$$

Equations (4.5) define a function which assigns a velocity (\dot{x}, \dot{y}) to every point (x, y) on the image plane to which a point in the scene is projected. This is the instantaneous image velocity field $\mathbf{v}(x, y)$.

From inspection of the right hand sides of (4.5) it is apparent that the velocity field has a component that depends on the translational motion of the camera (in plain brackets), and a component that depends on the rotational motion of the camera (in square brackets). Equations (4.5) could, therefore, have been derived by obtaining an expression for the image velocity field generated by pure rotation (\mathbf{R}) and an expression for the field generated by pure translation (\mathbf{V}) and adding them together. Only the translational part contains information about the scene since only this part depends upon the distances (Z) of surface points from \mathcal{O} . The rotational part depends only on the image coordinates (x, y) and not at all on the

scene⁶. Following Longuet-Higgins & Prazdny's notation the translational part of the image velocity field will be indicated by the superscript T and the rotational part by the superscript R, i.e., $\mathbf{v}^T(x, y)$ and $\mathbf{v}^R(x, y)$ respectively.

Since only the component of the image velocity field due to translation of the observation point depends on the distances of surfaces in the scene it alone carries information about the three-dimensional structure of the scene. It corresponds to the projection of the velocity field representation of Gibson's optic flow onto the imaging surface (the optic velocity field), an observation also made by Cutting, 1986. The rotational part only contains information about how the camera or eye is (instantaneously) rotating relative to the scene. This component of the field can, in effect, be completely cancelled by an appropriate eye movement (Koenderink, 1986). The fact that only the translational component is exterospecific serves to define a second problem of information pick-up, that of obtaining the translational component from the image velocity field where it is mixed up with the rotational component. This problem is discussed further in the next chapter.

Information in the Translational Component. The translational component of the retinal velocity field is described by the terms in plain brackets of equations (4.5). As is shown by Longuet-Higgins & Prazdny the velocities of every point in this field are directed towards or away from a unique "vanishing point". This point is the projection onto the imaging surface of either the "focus of radial outflow (FRO)"⁷ or of inflow which were identified by Gibson as features of the optic flow—only one or other of them can be *imaged* at one time on a retina.

The FRO is the point on the image plane which is pierced by the instantaneous translational velocity vector \mathbf{V} of the eye or camera, so the FRO lies in the instantaneous direction of translation (Gibson, 1950; Gibson, Olum & Rosenblatt, 1955). Gibson noted that the FRO therefore specifies the (instantaneous) direction of translation and could be used both to find and to control the direction of travel thus constituting one kind of potentially exploitable information.

It is worth noting at this point that failure to adhere to a rigorous terminology

⁶ Thus one cannot obtain any information about the three-dimensionality of the scene by making eye-movements (cf., Koenderink, 1986).

⁷ Gibson (1950, 1966) also referred to the focus of radial outflow as the *focus of expansion*. However, this is technically a misnomer since the expansion actually vanishes at this point (Koenderink & van Doorn, 1981), thus as Warren et al. (1988) note, Gibson's later (1979) name for it, which is given here, is to be preferred.

has created confusion over this issue in the past (e.g., Regan & Beverley, 1982). The problem arises because the retinal field will generally contain a rotational component which obscures the FRO. In fact, if an animal is looking at an object not along the direction of travel and the object's image is stabilized on the foveal region of the retina, the image velocity of points on the object projecting along the instantaneous line of sight will vanish. Effectively the retinal velocity field in the neighbourhood of the direction of gaze will have a form similar to that of an FRO (Regan & Beverley, 1982). This fact, which may be useful since it means that the retinal velocities of points on an interesting object's image are minimized, prompted Regan & Beverly to question the utility of the FRO. If looking at an object induces a singularity in the field which "looks" like an FRO, then how can an animal know that it isn't moving in the direction it is looking assuming it uses the FRO as information about the direction of travel? Clearly this is a question that would never had been asked if a rigorous terminology had been agreed upon. The FRO is a feature of the translational component of the image velocity field (Lee & Young, 1985; Torrey, 1985), and *a fortiori* of the optic flow, not of the image velocity field itself, except in the special case of the direction of gaze coinciding with the direction of motion.

In figure 4.2 the image velocity field due to instantaneous pure translation along the Z-direction perpendicularly towards a planar surface is shown. The FRO is at the centre and the velocity vectors increase with increasing distance from this point. The equation describing image point velocities in this case can be obtained from equations (4.5) by noting that $\mathbf{V} = (0, 0, V_z)$ and $\mathbf{R} = (0, 0, 0)$, we thus obtain from (4.5)

$$\dot{x} = \frac{xV_z}{Z}, \quad \dot{y} = \frac{yV_z}{Z}. \quad (4.6)$$

In the case of translation perpendicular to a flat surface (fig. 4.3), the distance, Z , is the same for every point on the surface. The image velocity (\dot{x}, \dot{y}) thus depends on the instantaneous velocity of translation, V_z , and the position (x, y) on the imaging surface—the further the image point (x, y) from $(0, 0)$ (the FRO) the greater its velocity. It is clear, on the other hand, that if the surface being approached is tilted relative to the direction of travel or not flat but lumpy and/or curved then the instantaneous distance Z will vary over the surface. Since the image velocity depends upon the distance Z of the surface point imaged (equations (4.6)) a curved, lumpy or tilted surface will impose additional structure on the translational velocity

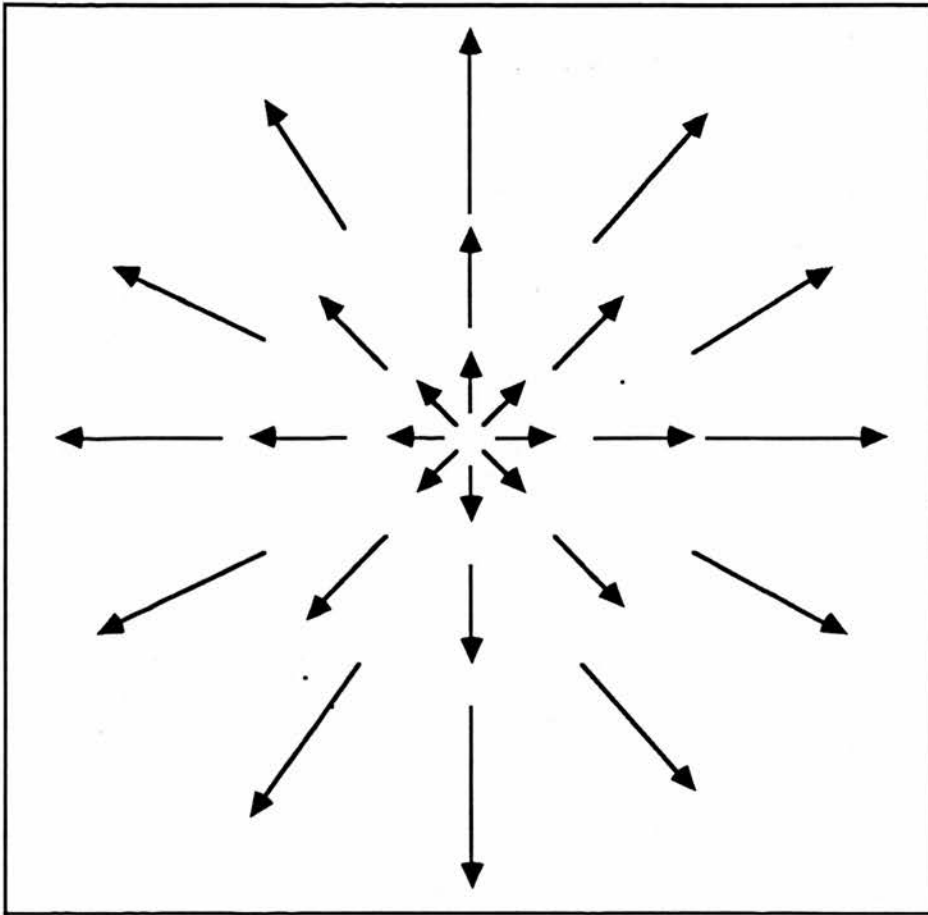


Figure 4.2. Image velocity vectors of selected texture elements on a plane surface being approached directly by an imaging system. The point of projection is heading for the point which is imaged at the centre (the FRO) and from which all the vectors appear to originate.

field as will separate surfaces lying at different depths.

Thus, the three-dimensional geometry of the scene imparts a geometrical structure to the translational field. On this basis it might be hoped that the translational field contains information about the depths, tilts or slopes and curvatures of surfaces in the scene. Indeed, mathematical analyses have shown that the depths of surface points scaled by the velocity of translation are specified (e.g., Gibson, Olum & Rosenblatt, 1955; Lee, 1974, 1980a; Longuet-Higgins & Prazdny, 1980), a fact that will be employed below where timing information is discussed.

It has also been shown that the slants (the orientation of a surface patch with respect to a visual direction) and distance scaled curvatures of smooth surfaces or surface patches are specified (e.g., Clocksin, 1980; Koenderink, 1985; Koenderink & van Doorn, 1975; Longuet-Higgins & Prazdny, 1980; Subbarao, 1988; Waxman

& Wohn, 1988). The mathematical analyses that have been used to demonstrate the presence of such surface information are local (i.e., restricted to a small neighborhoods of the field) and involve the spatial derivatives of the field (i.e., how fast the field is changing spatially—its gradient). Local spatial changes in the field generated by smooth surfaces reflect smooth changes in the depth of points on the surface due to the local slant or curvature of the surface. The mathematics that has been applied can be considered a power series expansion of the velocity field about a line of sight (Longuet-Higgins, 1986; Koenderink, 1986). There are two ways the analysis has been done; either in a coordinate based form where the field on a planar imaging surface is expanded about the Z-axis (Longuet-Higgins & Prazdny, 1980; Subbarao, 1988), or in a coordinate free (invariant) form over the manifold of visual directions (Koenderink & van Doorn, 1975; Koenderink, 1985, 1986). The slants of surface patches appear in the first order terms of the expansion (the gradient of the field), the scaled curvatures in the second order terms (Koenderink, 1986; Subbarao, 1988). Discussion of these methods of analysis and of the information they describe is beyond the scope of the present discussion. It should be noted in this context, however, that doubts have been raised as to the computational realism of obtaining information from spatial derivatives of the velocity field, especially the second derivatives of the field (Longuet-Higgins, 1986; Nakayama, 1985).

Finally, discontinuities in the field can be used to identify the boundaries (edges) of objects and surfaces (e.g., Clocksin, 1980; Koenderink & van Doorn, 1977; Nakayama & Loomis, 1974). The discussion which follows, however, is restricted to analysis of the information for timing interactions with an environment in motion relative to the actor. In summary, as noted by Gibson there are two kinds of information available in the translational component of visual flow: information about the relative motion of the observer (camera/eye) and the environment, which has been called *expropriospecific* (Lee, 1978); and information about the 3-D geometrical structure of the scene, *exterospecific* information.

§4.4. TIMING INFORMATION DURING LOCOMOTION

As was discussed in the previous chapter, an animal in motion through its environment requires information to correctly time acts such as jumping over something, landing on the ground, and placing the feet during running over stepping stones or rough terrain. In the present section it will be shown how information which could be used for the purpose of timing activities such as those just mentioned is present in the visual stimulus of an animal moving through a rigid environment. Specifically we shall be concerned with information present in the translational component of the image velocity field—the case that has been considered by Lee (1974, 1976, 1980a).

We have seen that the translational field contains an FRO, the retinal position of which depends on the direction of instantaneous translational motion. The retinal coordinates of the FRO are (x_0, y_0) ; at this point the image velocity of the translational field vanishes, $\dot{x} = 0$, $\dot{y} = 0$. Hence from the translational part of equations 4.5 ($R=0$) we obtain, at the FRO,

$$(x_0 V_Z - V_X)/Z = 0 \quad \text{and} \quad (y_0 V_Z - V_Y)/Z = 0, \quad (4.7)$$

which means that the coordinates of the FRO are

$$(x_0, y_0) = (V_X/V_Z, V_Y/V_Z). \quad (4.8)$$

Given (4.8) it is trivial to obtain the following relations (e.g., Longuet-Higgins & Prazdny, 1980):

$$Z/V_Z = (x - x_0)/\dot{x} = (y - y_0)/\dot{y}, \quad (4.9)$$

which demonstrate that the (instantaneous) velocity scaled depths (Z/V_Z) of every point in the scene imaged on the retina are available in the translational field, provided that the retinal coordinates of the FRO are known. If a locomoting animal knows its speed relative to the environment it can obtain the depths of all visible points in the scene from the translational field. Lee (1980b) has suggested that a legged animal might be able to estimate its stride length and its stride time and thereby have access to an estimate of its velocity since this will be equal to the stride length divided by the stride time. Without such velocity information the *relative* depths of all imaged points are specified during motion through a rigid environment. If two image points (x_1, y_1) and (x_2, y_2) are at depths Z_1 and Z_2

respectively, then the relative depth Z_1/Z_2 can be obtained by division of equations (4.9) defined for the two points yielding, for the x - coordinate,

$$\frac{Z_1}{Z_2} = \frac{(x_1 - x_0)\dot{x}_2}{(x_2 - x_0)\dot{x}_1} \quad (4.10)$$

and similarly for the y coordinate. The 'absolute' depths are not specified, which is only to be expected since absolute depth implies a scale of measurement, i.e., some fixed 'standard' distance in terms of which any depth can be represented. If some (known) fixed distance is optically available it is possible to scale the depth information. An animal's eye-height has been suggested as a possible standard (Lee, 1974; Sedgewick, 1973). However, even without a standard distance for scaling depth the quantity Z/V_z is potentially useful for the control of action since it has the dimensions of time and represents the depth in *temporal* terms. The representation of the translational velocity field given in equation (4.9) can be thought of as an instantaneous temporal depth map of the scene, in which points in the scene are represented not in terms of their nearnesses in space but in terms of their instantaneous nearnesses in time.

As stressed by Lee (1980a,b), if the Z -component of translational velocity (V_z) remains constant, the velocity scaled depth is the time remaining before the environmental point reaches the X - Y plane (see fig. 4.2)—the *time-to-contact*⁸ with this plane. Contact with this plane when there is no actual contact with the moving animal is also called the "time-to-nearest approach" by Lee & Young (1985). Thus a locomoting animal has access to information at every 'instant' (up to a visuo-motor delay) about the time remaining before a point in the scene will contact the X - Y plane through the centre of projection of its eye assuming the velocity remains constant.

The Variable Tau. Lee (1976, 1980a,b) introduced the symbol τ (tau) to represent the image plane variable (or optic variable) that specifies the velocity scaled depth or time-to-contact (T_c). From equations (4.9),

$$\tau = (x - x_0)/\dot{x} = (y - y_0)/\dot{y}. \quad (4.11)$$

This differs slightly from the definition of τ that appears in Lee (1980a,b) where $\tau = r/\dot{r}$, with r and \dot{r} defined as in figure 4.3; the two definitions are equivalent.

⁸ Time-to-contact will be abbreviated as T_c in what follows.

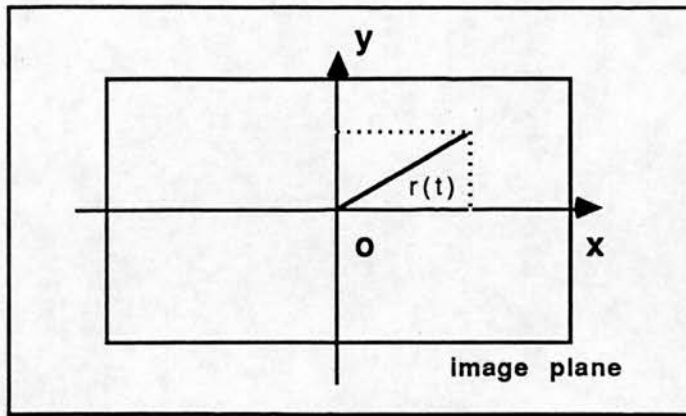


Figure 4.3. Lee's (1980a) definition of tau places the FRO at the origin of the image plane coordinate system (o). $r(t)$ is the magnitude of the position vector of a visible feature on the image plane and the velocity of the feature is directed away from o having magnitude $\dot{r}(t)$.

Tau is here defined with reference to a feature of the translational image velocity field, the FRO. It is possible, however, to define tau independently of the FRO as Lee (1976) has done. Here we will give an illustration of how this can be done, later it will be done in fuller generality using spherical rather than planar coordinates. Consider a perpendicular approach to a plane textured surface — this surface is part of a rigid, fixed environment through which the observation point is moving. Suppose that there are two points on this surface whose x coordinates on the image plane are x_1 and x_2 , figure 4.4, the y coordinates could be considered instead, or as well, but they will be ignored here since our purposes are primarily illustrative. From equations (4.9) it is a simple matter⁹ to obtain

$$\tau = \frac{x_2 - x_1}{\dot{x}_2 - \dot{x}_1}. \quad (4.12)$$

The left hand side of (4.12) is the reciprocal of the relative rate of separation of the two image points along the x direction and it specifies the T_c of the plane surface with the X-Y plane (see footnote 9). Tau considered in this way can be given a coordinate free definition as will be shown below. There is thus a distinction between τ defined using the FRO and τ defined without using any feature of a translational velocity field. The two will be distinguished in the following way: τ defined using the FRO will be called *global tau* and denoted τ_g , τ defined at a locality in the image will be called *local tau* and denoted τ_L .

Tau under Spherical Projection. It so happens that the expressions derived so far using a plane projection surface can be uniquely transformed into corresponding expressions on a spherical surface as mentioned earlier (cf Lee, 1974; Longuet-Higgins

⁹ $Z/V_Z = x_1/\dot{x}_1 = x_2/\dot{x}_2$ can be written down directly from (4.9) and $x_1\dot{x}_2 = x_2\dot{x}_1$ obtained. Subtracting $x_1\dot{x}_1$ from both sides of the latter one gets, $x_1\dot{x}_2 - x_1\dot{x}_1 = x_2\dot{x}_1 - x_1\dot{x}_1$ which can be rearranged to give $(x_2 - x_1)/(\dot{x}_2 - \dot{x}_1) = (x_1/\dot{x}_1)$, from which (4.12) is immediate.

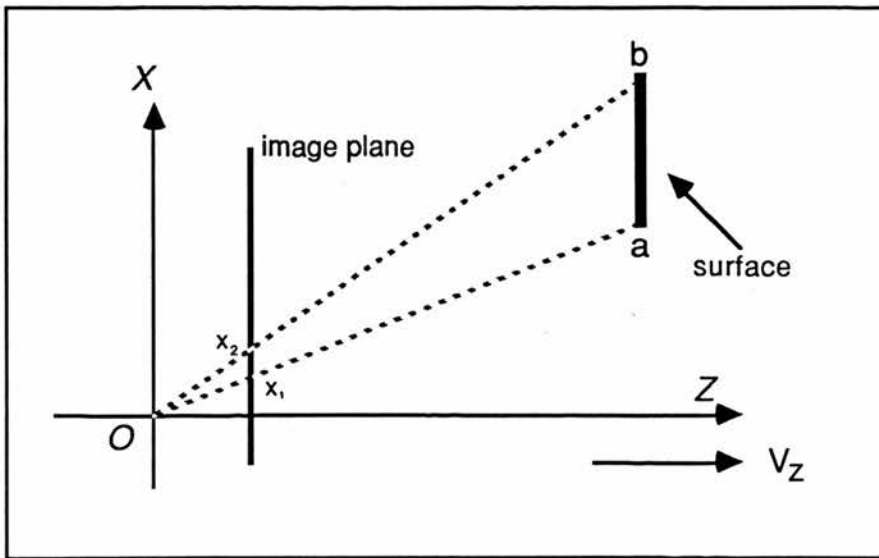


Figure 4.4. Two coplanar points a and b on a surface in the scene are imaged at points x_1 and x_2 on the image plane. The imaging system is moving along the Z direction with speed V_z .

& Prazdny, 1980). Later on it will be convenient to use spherical projection and for this reason we will transform the expressions for T_c information that have been obtained above for a plane surface into corresponding expressions on a spherical surface in this section. The problem can be viewed as simply one of transforming from Cartesian coordinates in the image plane (x, y) to spherical polar coordinates (θ, ϕ). As mentioned above the spherical coordinates serve as labels for the visual directions.

Let us consider, as above, only the translational component of the velocity field over the set (manifold) of visual directions. If the FRO has instantaneous spherical coordinates (θ_0, ϕ_0) then by noting that the transformation between the Cartesian coordinates and the spherical coordinates is given by $x = r \tan \theta$ and $y = r \tan \phi$ (with the distance r from the observation point equal to one), relations (4.9) can be transformed into angular form to give

$$Z/V_z = (\tan \theta - \tan \theta_0)/\dot{\theta} \sec^2 \theta = (\tan \phi - \tan \phi_0)/\dot{\phi} \sec^2 \phi. \quad (4.13)$$

If the FRO is at the point $(0,0)$, (4.13) reduces to

$$Z/V_z = \frac{\sin 2\theta}{2\dot{\theta}} = \frac{\sin 2\phi}{2\dot{\phi}}, \quad (4.14)$$

This shows that the instantaneous time-to-contact with the plane normal to the direction of translation is specified in terms of variables of the angular velocity

field. When θ and ψ are small (4.14) reduces to

$$Z/V_z \approx \theta/\dot{\theta} \approx \phi/\dot{\phi}. \quad (4.15)$$

The quantities on the right hand sides of (4.14), (4.14) and (4.15) are angular versions of quantities that were called τ_g in the previous section and may, therefore, also be called τ_g for consistency (cf. Lee & Young, 1985).

Local tau defined on a spherical projection surface is defined by Lee in two slightly different ways. Firstly, he defines it in terms of two image points, corresponding to how it was defined in equation (4.12), (Lee, 1976; Lee & Young, 1985). If the angle subtended by two points on an object (as in figure 4.4) at the point of observation is α , then (local) tau (τ_L) is defined in Lee & Young (1985) to be

$$\begin{aligned} \tau_L &= \frac{\cos \alpha \sin \alpha}{\dot{\alpha}} \\ &= \frac{\sin 2\alpha}{2\dot{\alpha}}. \end{aligned} \quad (4.16)$$

Typically the angle α will be small and hence τ_L will be given by $\alpha/\dot{\alpha}$ (this is how Lee (1976) defines tau). The second definition involves the area of the image of an optic texture element or of an object subtending a small visual angle. Lee & Young (1985) define local tau to be (twice) the inverse of the relative rate of dilation of the image of an object or texture element; if the area of the image on a spherical projection surface (a unit distance from the projection point) is Ω then Lee and Young's definition reads

$$\tau_L = 2\Omega/\dot{\Omega}. \quad (4.17)$$

This is the usual way in which tau is understood (see e.g., Turvey & Carello, 1986). Defined in either of these two ways, local tau specifies the T_c of the texture element or object with the point of observation in the case when this point is on a direct collision course with the object or texture element.

Equations 4.16 and 4.17 represent coordinate free definition of tau — no origin or set of coordinate axes on the imaging surface are required to measure τ_L defined in these ways, one needs only a scale of angular measurement. Note that τ_L so defined only provides time to contact information if the target object is not rotating, or if it is, it is spherical. As mentioned in the introduction to this chapter, the consequences of not fulfilling these conditions are examined in chapter 5.

§4.5. INFORMATION ABOUT MOVING OBJECTS

In §3 and §4 the image velocity field due to observer motion relative to a fixed and rigid environment was considered. Typically, however, there will be objects other than the observer moving in the environment. Information about the motion of such objects relative to the observer is required for many actions; for example, a player catching and hitting balls in games such as tennis, cricket and baseball or an animal in pursuit of moving prey. As was pointed out in chapter 3, to time a catch or a hit, the time-to-contact of the object to be intercepted with the point at which interception will take place is required. The question to ask, therefore, is whether such information is available in the visual stimulus.

First of all it will be useful to establish that local tau for a target object specifies the same environmental quantities when both the target and the observer are in motion as it does when only the observer is moving. We continue to assume, of course, that the target satisfies the conditions assumed throughout this chapter and stated in the introduction which are required for local tau to provide useful timing information. Referring back to figure 4.1, suppose that the point \mathcal{P} now lies on an object moving through the environment with an instantaneous velocity \mathbf{U} . The velocity $\dot{\mathbf{P}}$ of \mathcal{P} in the moving coordinate frame is now given by $\dot{\mathbf{P}} = \mathbf{U} - (\mathbf{V} + \mathbf{R} \times \mathbf{P})$. If the instantaneous components of \mathbf{U} along the X, Y and Z directions are U_X, U_Y and U_Z then the relative *translational* velocity of the object relative to the \mathcal{OXYZ} frame will be a new vector \mathbf{W} where $(W_X, W_Y, W_Z) = (U_X - V_X, U_Y - V_Y, U_Z - V_Z)$. The expression for the overall relative velocity of object and frame is identical to equations 4.2a-c save that where a component of \mathbf{V} occurs in equations 4.2 the corresponding component of $-\mathbf{W}$ appears in its stead. This means that an expression for the local velocity field over the image of the moving target object can be obtained. This field has a translational component due to motion of the object and motion of the observer and superimposed on this will be a rotational field due to rotation of the \mathcal{OXYZ} coordinate frame (object motion does not contribute to the rotational field). As might be expected, only the translational component is different from the case when the target object is stationary; the relevant expression for the translational component of the field is

$$\begin{aligned}\dot{x} &= (-xW_Z + W_X)/Z \\ \dot{y} &= (-yW_Z + W_Y)/Z.\end{aligned}\tag{4.18}$$

Is time-to-contact information available from the combined field in the same way as from the field due to pure translation of the eye? Global tau values for points on the object will not provide useful timing information since the global rigidity constraint required for the definition of global tau is violated here. What about

local tau? Consider two points on the object that is being considered which project to two points on the image plane with image coordinates (x_1, y_1) and (x_2, y_2) respectively. From equation 4.18 the velocity of the point (x_1, y_1) is given by

$$\begin{aligned}\dot{x}_1 &= [-x_1 W_Z + W_X]/Z, \\ \dot{y}_1 &= [-y_1 W_Z + W_Y]/Z\end{aligned}$$

and similarly for the point (x_2, y_2) . If we take the velocities in the x direction only we obtain two equations from which W_X can be eliminated to give

$$\frac{-Z}{W_Z} = \frac{x_2 - x_1}{\dot{x}_2 - \dot{x}_1}. \quad (4.19)$$

The right hand side of this equation is identical to the right hand side of equation 4.12 and is thus local tau for the two points. The left hand side of the equation is the time-to-contact of the object with the X-Y plane. This shows that when both the target object and observer are moving provided the usual conditions are met, local tau specifies the same time-to-contact as when the target is stationary. This is what is expected; the expansion of the target object depends only upon the relative translational motion of the observer and target provided the conditions required for defining timing information using local tau are satisfied.

More General Timing Information The T_c information that has been described so far is limited to giving time-to-collision with the observation point or the time-to-nearest approach. It would be desirable to show that information about the time-to-contact with any given interception point is available.

It is most convenient to use a spherical projection surface to examine this question. The situation that will be dealt with is illustrated in figure 4.5. The moving object is spherical, e.g., a ball, and is considered to be moving at constant speed along a straight line path. Lee & Young (1985) have considered this case and shown how what they call the "time-to-nearest approach" (T_n) is specified. The point of nearest approach is the point on the path of the object which is closest to \mathcal{O} (n in fig. 4.6); the line joining \mathcal{O} to n is perpendicular to the object's path. To show that T_n is specified by variables of projection of the object's motion on the spherical surface it suffices to differentiate the geometrical relations $\sin \theta = D/S$ and $\cos \theta = L/S$ derived from fig. 4.5. Making the appropriate substitutions one obtains the expression

$$1 + \left(\frac{S}{\dot{S}}\right)^2 \dot{\theta}^2 = \frac{\dot{D}S}{D\dot{S}} \quad (4.20)$$

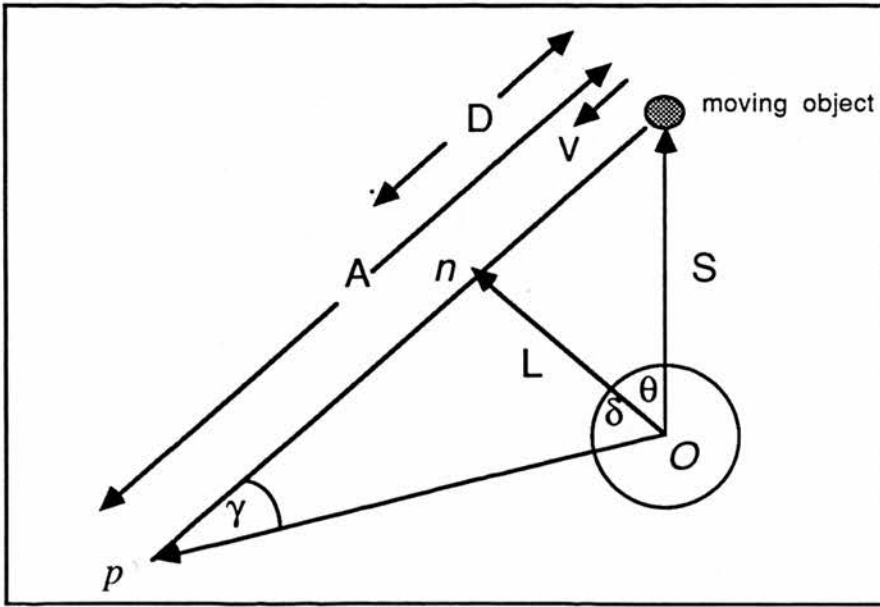


Figure 4.5. Geometry of time-to-contact. A spherical object (e.g., a ball) is moving in the plane of the page towards the point p past the centre of projection O . The point at which it passes closest to O is marked n . The angle ψ in the text is equal to $\theta + \delta$. The velocity of the object is considered constant. The circle represents a slice through a spherical imaging surface centred on O .

Here $-D/\dot{D}$ is the time to nearest approach and the quantity $\dot{\theta}$ has the same value as the rate of change of any angle subtended at O between the moving object and a point on its path in the fixed environment.

It is the rate of change of direction of the object relative to O and will be denoted u after Lee & Young (1985). The optic variable τ_L is the quantity $2\Omega/\dot{\Omega}$ where Ω is the area of the ball's image (Lee & Young, 1985) and is equal to $-S/\dot{S}$ which follows from the discussion in §4. Substitution of τ_L and u into (4.18) yields

$$T_n = \frac{\tau_L}{1 + \tau_L^2 u^2}, \quad (4.21)$$

which is Lee & Young's equation (2). Referring to figure 4.5 again, suppose that the point p is any point on the future path of the moving object. Application of the sine rule yields

$$\frac{\sin \gamma}{S} = \frac{\sin \psi}{A}$$

differentiating this with respect to time, substituting for $\sin \psi$ and rearranging one obtains

$$(\dot{A}/A) \sin \psi = (\dot{S}/S) \sin \psi + \dot{\psi} \cos \psi, \quad (4.22)$$

since the angle γ is constant. Rearranging (4.22) and substituting τ_L for $-S/\dot{S}$ gives

$$\frac{A}{\dot{A}} = \frac{\tau_L}{\tau_L \dot{\psi} \cot \psi - 1},$$

measuring quantities from the fixed point p and noting that the quantity $-A/\dot{A}$ is the time-to-contact of the object with the point p , T_p , and that $u = -\dot{\psi}$ we can write the above equation as

$$T_p = \frac{\tau_L}{1 + \tau_L u \cot \psi}. \quad (4.23)$$

(4.23) shows that at every instant of time the time-to-contact, T_p , of a moving object with a point p anywhere on its future path is optically specified provided the object's velocity remains constant and that the *direction* of the point p is optically available. The direction of p could be the same as that of some point in the fixed environment. This environmental point could then serve as a label for the direction of p . An observer wishing to intercept the object is then able to do so along any direction from \mathcal{O} through which the object will pass (intersecting the direction at a point p) provided a point (texture element) in the fixed environment lies in that direction. In the next chapter we will see how this restriction can be lifted.

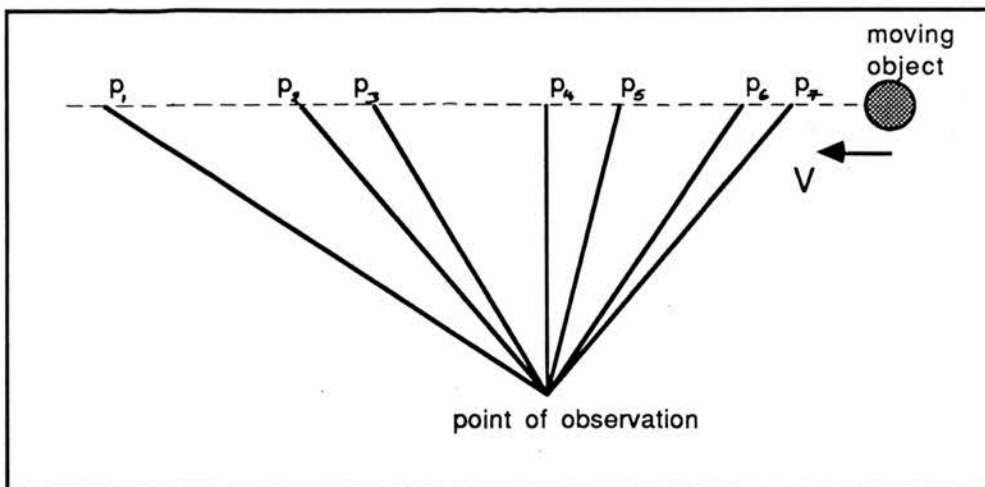


Figure 4.6. An object moving with constant velocity in the plane of the page will pass through a sequence of points (some of which are marked p_i these can be thought of as the directions from the observation points of visible features in the fixed environment) each in a unique visual direction.

ches

A further observation can be made regarding the information described in equation 4.23. Consider the situation illustrated in figure 4.6. At every instant of time a T_c with each point p_i on the object's path can be defined. Each point is in a particular visual direction making an instantaneous angle ψ_i with the instantaneous direction of the moving object. Thus, for any point p_i in a labelled visual direction the visual quantity $\tau_L/(1 + \tau_L u \cot \psi_i)$ is perceptually available. A suitably equipped visual system is thus, in principle, able to compute the T_c with any labelled visual direction. It may be possible to extract, in parallel, all such time-to-contacts at once. Such a process would supply a T_c value for every upcoming point on the object's path lying in a labelled visual direction. This is exactly the kind of information which we suggested in the last chapter would be convenient in acts such as catching and hitting. This issue is discussed further in the next chapter.

Timing Information Without Local Tau Consider either a moving object heading for a collision with some other object or surface a long way from the observer or a "ball" moving towards a "bat" in a game of video tennis. In both these cases the two 'objects' in collision are known to lie on a plane which is perpendicular to the line of sight and hence to parallel to the image plane (in the first case the great distance from the observation point guarantees that even if the two objects do not both lie on such a plane it is likely to be a good approximation to assume that they do).

Coplanarity means that we can apply the optic geometry diagrammed in figure 4.7. From this geometry it is simple to derive the following relation (see caption to figure 4.7) which shows how the time remaining before the two objects collide is visually available,

$$T_c = w/\dot{w}. \quad (4.24)$$

This is rather obvious but can be used in the derivation of a more general relation giving potentially the same information as equation 4.23.

Consider the geometry diagrammed in figure 4.8, the following relation can be obtained by application of the sine rule,

$$D = W \sin \varphi / \sin \vartheta, \quad (4.25)$$

differentiating (4.25) with respect to time one obtains

$$V = \frac{\sin \vartheta (\dot{W} \sin \varphi + W \dot{\varphi} \cos \varphi) - W \dot{\vartheta} \sin \varphi \cos \vartheta}{\sin^2 \vartheta}. \quad (4.26)$$

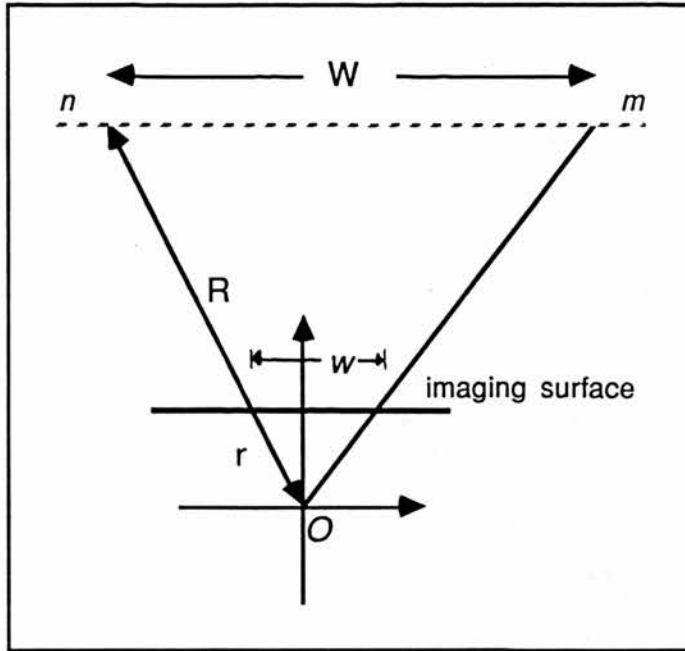


Figure 4.7. Instantaneous geometry of time-to-contact in the video-game case. Two collinear points m and n project onto a slice through a planar imaging surface. If the point n is considered motionless and the other as moving towards it in the plane of the figure, then by similar triangles we have $W/R = w/r$ which when differentiated with respect to time gives, $\dot{W}/R = \dot{w}/r$. Division of these two relationships gives equation 4.24 in the text (where $T_c = W/\dot{W}$). Note that this argument is unchanged when points n and m are both moving, as is readily verified.

Dividing (4.26) by (4.25) yields

$$\frac{V}{D} = \frac{\dot{W}}{W} + \dot{\varphi} \cot \varphi - \dot{\vartheta} \cot \vartheta. \quad (4.27)$$

Noting that V/D is the reciprocal of the time-to-contact T_p , that $\dot{W}/W = \dot{w}/w$ and $\vartheta = 180 - \varphi - \beta$ where β is constant, from (4.27) one obtains

$$T_p^{-1} = (\dot{w}/w) + \dot{\varphi}[\cot \varphi - \cot(\varphi + \beta)]. \quad (4.28)$$

When $\beta = 0$ (4.28) reduces to (4.24) as expected. All the variables appearing on the right hand side of equation (4.28) are potentially available from the projection except the angle (β) that the velocity vector makes with the image plane. How can this angle be obtained? Lee & Young (1985) discuss a monocular method for obtaining the direction of the velocity vector which involves the use of τ_L (see also Todd, 1981, for a similar method). If such a method is used then equation (4.28) has no advantages over equation (4.23) and is considerably more long-winded. It is

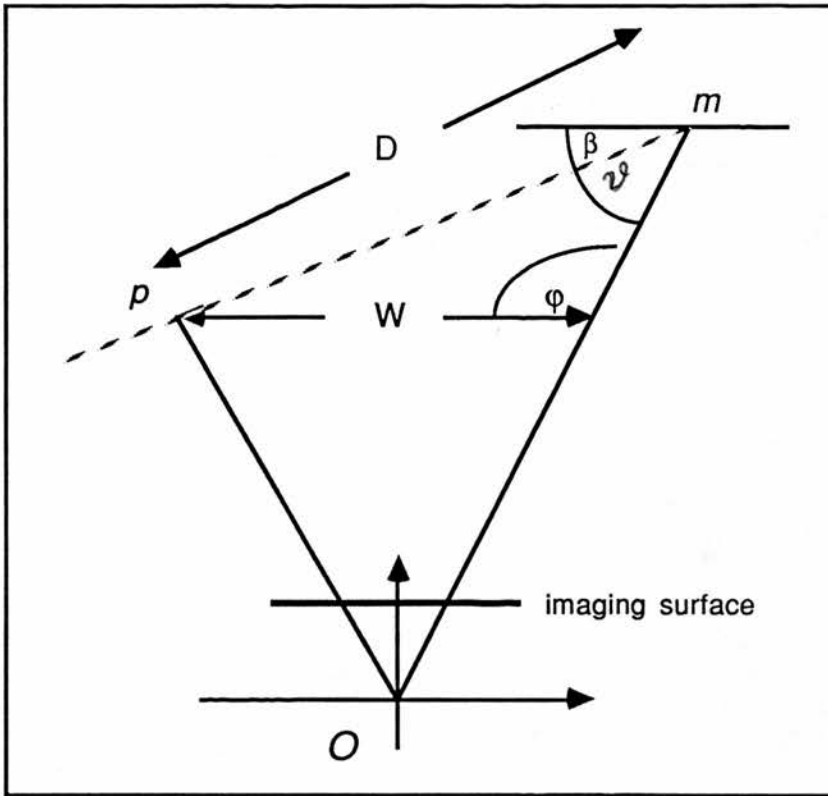


Figure 4.8. Instantaneous geometry of time-to-contact for arbitrary planar motion. A point m is moving along the path shown as a dotted line with constant velocity (speed V) making an angle β with the image plane. The angular quantities can be transformed into variables of planar projection but the corresponding expressions for time-to-contact are more complicated, for the sake of clarity angular quantities are represented as such in the text.

possible, however, that a binocular observer can obtain the direction of the velocity in a manner independent of τ_L by making use of the binocular flow information about the direction of motion in depth described by Regan and Beverley (1979).

That the information is binocularly available can be demonstrated by introducing a second imaging system into figure 4.8 slightly to the left, say, of the system illustrated and having its image plane oriented parallel to that of the illustrated system. We can clearly write the following two equations for these two imaging systems indicating that variables are defined for the left hand system by the subscript L and for the right hand system by the subscript R ;

$$T_p^{-1} = (\dot{w}_R/w_R) + \dot{\varphi}_R[\cot \varphi_R - \cot(\varphi_R + \beta)]$$

and

$$T_p^{-1} = (\dot{w}_L/w_L) + \dot{\varphi}_L[\cot \varphi_L - \cot(\varphi_L + \beta)].$$

We now have two equations in two unknowns (T_p and β) hence T_p is visually

specified without requiring local tau. Note that equation 4.24 provides a good approximation to T_p when w is small and may therefore be used instead of equations 4.29 or 4.23 in some circumstances. What information observers actually use is a matter for empirical investigation.

§4.6. CONCLUSIONS

In this chapter a analysis of the timing information available in the visual stimulus was presented. All the information described relied upon two assumptions:

- That the environment as a whole was rigid (for the definition of global tau) or that the moving object was rigid (for the definition of local tau).
- That the velocity field could be extracted from the visual flows constituting the primary input (in the case of global tau we required further that the translational component could be extracted).
- That the object or surface for which local tau is defined is spherical or lies normal to the line of sight and does not rotate.

There was a further assumption made in the definition of local tau which has important consequences for the extraction of this information. This assumption is made explicit and discussed at length in the next chapter. Provided that the requirements of all the assumptions can be met then there is a good deal of potentially available timing information in the visual input. In addition, throughout the analysis relative motion between the observer and the environment was considered to have constant velocity. This is reasonable if the constant velocity strategy is used by humans engaged in interceptive actions (see chapter 3). It is possible to derive expressions for time-to-contact without the assumption of constant velocity and some simple cases are dealt with briefly in Appendix A.

It will perhaps be instructive to express in a different form the timing information described in this chapter. First of all tau can be generalised to mean any inverse of a rate of dilation of a sensory quantity (cf Lee, 1990). Thus the quantity w/\dot{w} appearing in equation 28 may be called tau. Secondly, the following two quantities may be introduced: the reciprocal of the time-to-contact or *immediacy* which will be symbolised as γ (Koenderink, 1985), and the reciprocal of tau (the rate of dilation) which may be symbolised¹⁰ as μ . To re-express the various

¹⁰ The symbol μ is used here after Longuet-Higgins & Prazdny (1980) who use it to symbolise the divergence of the velocity field — the rate of dilation of a small region of the field may be used to approximate the divergence.

times-to-contact that were derived in this chapter in terms of the concepts just introduced, the following notation will be used: γ – the immediacy of contact with the image plane or with the observation point; γ_n – the immediacy of nearest approach; γ_p – the immediacy of contact with some point p on the upcoming path of a moving object; μ_g – the reciprocal of global tau; μ_L – the reciprocal of local tau; μ – the reciprocal of tau defined neither globally nor locally. Using this notation we may write:

$$\begin{aligned}\gamma &= \mu_g, \\ \gamma_n &= \mu_L + \mu_L u^2, \\ \gamma_p &= \mu_L + u \cot \psi, \\ \gamma_p &= \mu + \dot{\varphi} [\cot \varphi - \cot(\varphi + \beta)].\end{aligned}$$

These equations are re-expressions of equations 4.8, 4.21, 4.23 and 4.28, respectively. Expressed in this form it is easy to see that the nearness in time (the immediacy) is always given by a rate of dilation (μ) plus “something else”.

In the next chapter some of the problems that must be solved if the information described in this chapter is to be extracted (picked-up) and used are discussed.

5

Obtaining Timing Information

§5.1. INTRODUCTION

In the previous chapter some of the information that is potentially available in the translational component of the image velocity field was described. As indicated, there are certain problems involved in extracting this information since it is not *explicit* in the image flow. The first problem in making the information explicit is actually measuring motion in the image. This problem, though important, will not be discussed further since it is not of direct concern to us here and in any case is reviewed extensively elsewhere (e.g., Marr, 1982; van Santen & Sperling, 1985; Hildreth & Koch, 1987). The second problem is the extraction of the image velocity field from the measured visual motion, sometimes called the *visual motion field* (e.g., Verri & Poggio, 1986); the fact that these two fields do not, in general, coincide has important implications for the extraction of local tau. The third problem is extracting the translational component of the image velocity field which must be done before global tau can be extracted. This component may be obtained either by stabilizing the eye such that there is no rotational component (which would involve pointing the eye in the direction of motion and is, therefore, rather inflexible) or by filtering out the rotational component (Lawton, Reiger & Steenstrup, 1987; Longuet-Higgins & Prazdny, 1980; Prazdny, 1981). A final problem concerns whether visual processes alone can be considered responsible for extracting the information described in the previous chapter or whether other perceptual systems are involved. In this chapter these questions of information extraction are discussed in relation to the timing information described in chapter 4. In the section which follows, the problems that arise in interpreting local tau as time-to-contact information when the conditions mentioned in chapter 4 are not fulfilled are examined.

§5.2. PROBLEMS INTERPRETING LOCAL TAU

As described earlier (chapter 4, §2) the primary visual stimulus is a spatio-temporal pattern of light intensity flowing over the retinae. Spatial changes in image intensity will normally be due to different surfaces in the environment and different regions within such surfaces having different light reflectance properties and/or being in different conditions of illumination. Motion of objects in the environment and motion of the observer through the environment leads to motion of the spatial pattern of these intensity changes in the retinal image. However, motion of image intensity changes does not always correspond to the projection onto the imaging surface of motion of identifiable physical entities in the environment (e.g., physical points forming the edge of an object or forming the boundary of surface regions of different visible texture). It is easy to see that this is the case by considering situations in which the imaged environment contains shadow boundaries, deforming patterns of shading, “self-occluding” object boundaries (defined below) and specular highlights. These phenomena give rise to spatial changes in image intensity whose motion will not, in general, correspond to the motion of any objects or surfaces in the environment. Consider the following example, due to Todd (1985): a smooth solid object shaped like a rugby ball is imaged on a surface. The object is rotating about a vertical axis through its centre of mass such that at two different times it is imaged as an oval and as a circle as illustrated in the following diagram (figure 5.1). As the object rotates the image contour is in continuous motion—the shape it bounds on the imaging surface is deforming, from an oval into a circle and back. Such an image contour is often called a “self-occluding boundary” of the object (Marr, 1982; Todd, 1985). The physical points which correspond to the image contour are clearly not the same at any two different times. Indeed, as the object shown imaged in the figure rotates from its position at t_1 (figure 5.1a) to its position at t_2 (figure 5.1b) the image contour moves in one direction in the image plane (right in the figure), whereas a physical point on the object’s surface corresponding to a point on the image contour at t_1 (indicated by the filled circle) moves in the opposite direction (left).

These phenomena illustrate the fact, intimated in chapter 2, that the image velocity field does not, in general, correspond to the image motion field. The image velocity field is only *implicit* in the image motion field because the latter contains motion which does not correspond to the projected motion of identifiable physical

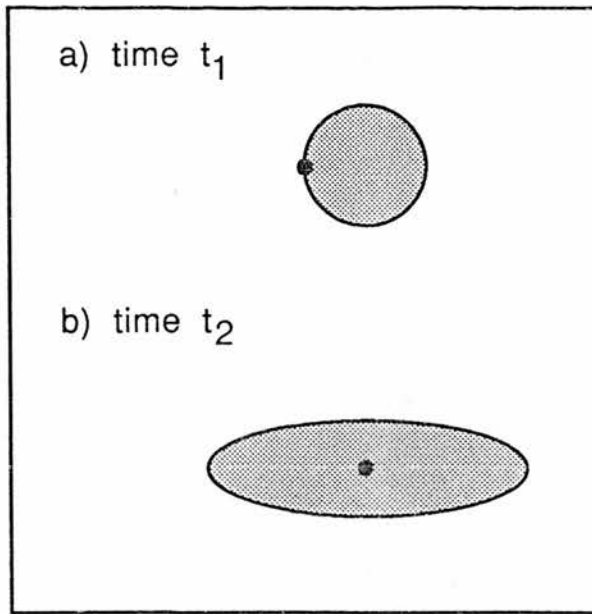


Figure 5.1. A rugby ball shaped object rotates about a vertical axis (parallel to the plane of the page) through its centre and is imaged on a plane as a circle (a) at time t_1 , and as an oval (b) at time t_2 . A physical point (black circle) which is imaged on the contour at t_1 , lies within the image contour at t_2 , as shown.

features (edges, texture elements, etc.). Some kind of perceptual processing is needed if the image velocity field is to be made *explicit*. In order to do this some process(es) is (are) required that can distinguish visual motion due to the projected motion of identifiable physical features from visual motion due to the phenomena described above. Both types of motion contain useful information: information in the image velocity field was extensively discussed in the last chapter and deforming optical contours due to motion of objects with self-occluding boundaries contain information about solid shape (Koenderink & van Doorn, 1976, 1986; Koenderink, 1985) which can be used by human observers (Todd, 1985).

The preceding considerations have implications for the usefulness of local tau in defining time-to-contact information. In the last chapter it was assumed that changes in the size of an object's image were due solely to changes in the distance of the object from the observer. In section 4.4 the moving object was considered to be a rigid sphere (e.g., a ball). This guarantees that any changes in the size of the object's image are due to changes in the distance of the object from the observation point. When the object is not spherical there can be changes in the

image size which are due to different regions of the surface of the object being visible as the direction of the object from the imaging system changes as it moves. If the object were stationary, similar changes in its image would be generated by rotating the object (clearly, rotating a spherical object does not change the size of its image). In Todd's displays involving objects with self-occluding boundaries the image size changed without any change in distance from the viewer.

The different types of object motion which influence image size when the rigidity constraint is satisfied are summarized in figures 5.1 and 5.2: figure 5.1 illustrates self-occlusion, figure 5.2a illustrates foreshortening and 5.2b illustrates dilation of texture due to rotation. These effects lead to changes in the size of the image of an object or bounded surface region which are not due to the target's approach to the observer, consequently local tau does not specify instantaneous time-to-contact with the observation point. Consider figure 5.2a, the solid angle, $\Omega(t)$, subtended at the observation point at time t (provided the angle is small) is given by

$$\Omega(t) = \frac{A \cos \theta(t)}{R^2(t)} \quad (5.1)$$

If θ is not constant, either because the object is rotating or because it is not directly approaching the observation point, the following expression for the relative rate of dilation of the solid angle is obtained:

$$\frac{\dot{\Omega}(t)}{\Omega(t)} = \frac{2V}{R(t)} - \dot{\theta}(t) \tan \theta(t). \quad (5.2)$$

The right hand side of this equation is the sum of two terms: the first depends on the approach of the target to the observation point and is twice the immediacy of contact with this point. The second term depends on θ and its rate of change with time. What effect does this second term have upon taking local tau as an estimate of the time-to-contact with the observation point? The error function is plotted in figure 5.3 at two different times-to-contact. It is clear from the figure that quite serious errors would result from using local tau as an estimate of time-to-contact.

In certain cases of the type shown in figures 5.1 and 5.2a where local tau does not give time-to-contact with the observation point, computation of time-to-contact is still possible. Consider figure 5.4, an object approaches a point of observation directly whilst rotating about one of the axes marked a and b . The relative rate of

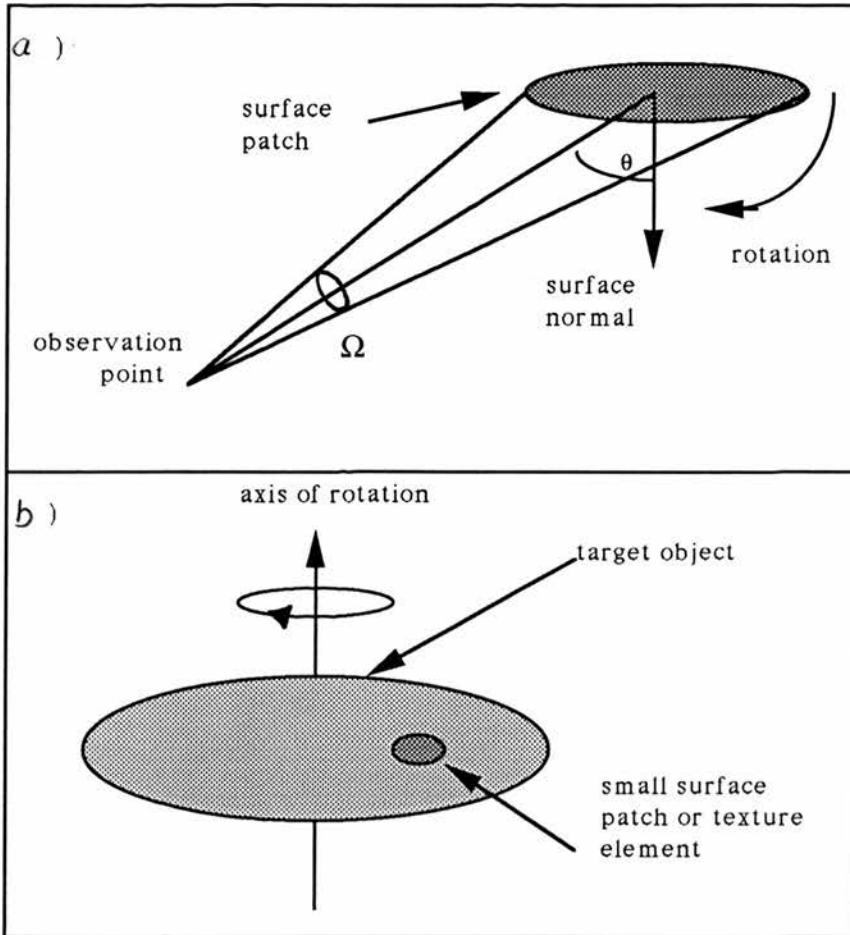


Figure 5.2. a) Foreshortening: a small surface patch or object such as a coin is rotating about an axis bisecting it and lying parallel to the plane of the page. As it rotates in the sense illustrated, the angle θ changes and hence the solid angle Ω subtended at the observation point changes. b) A small surface patch not lying along the projection of the axis of rotation onto the plane of the page will be instantaneously moving towards or away from the observation point due to the target objects' rotation. This will lead to changes in the angle subtended at the observation point by the surface patch and hence local tau for the patch will not give accurate time-to-contact information.

dilation of the solid angle subtended by the target object at the observation point is given by equation 5.2. However, consider the image of the object along the axis of rotation only. If the axis is b then the object along this axis (the line joining the two points p and q shown in the figure) subtends a visual angle γ . The reciprocal of the

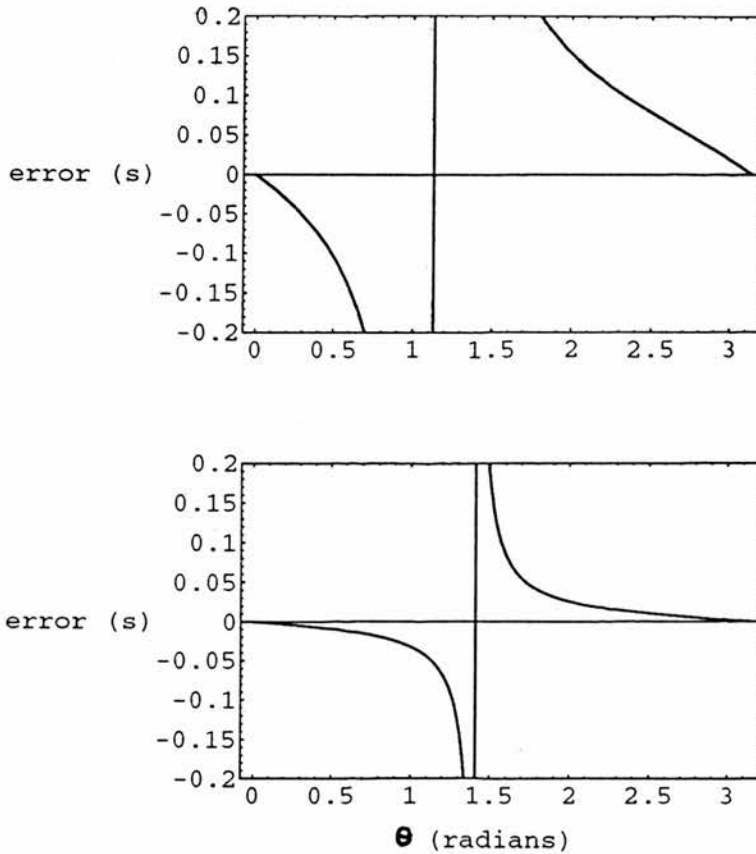


Figure 5.3. The difference between the actual time to contact and the estimate provided by local tau (the error) plotted as a function of θ when the target (as in fig. 5.2b) is rotating through half a revolution every second. Top, 300 milliseconds and bottom, 100 milliseconds before contact.

relative rate of dilation of γ is the time-to-contact of the target with the observation point (see, e.g., Lee, 1976). The quantity $\gamma/\dot{\gamma}$ (which corresponds to what Lee, 1976, called tau and here corresponds to a variety of local tau) can be measured from an imaging surface provided that the points p and q are not occluded by other parts of

the object during its rotation. If, however, rotation is about some other axis, such as that marked c in figure 5.4, then it is more difficult to obtain the desired time-to-contact. Measurement of local tau along a single dimension will no longer work because no two visibly identifiable points or features on the target (like p and q in the example just presented) lie on the axis of rotation's projection onto the imaging surface.

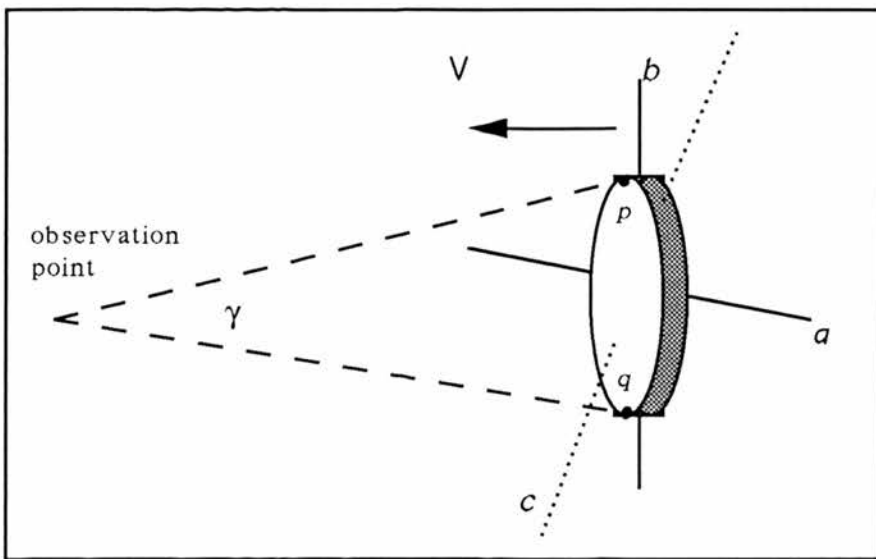


Figure 5.4. A target (which might be a small surface patch as in fig. 5.2b for example) approaches an observation point with velocity V . The lines a , b and c are three possible axes of rotation.

The above discussion may be summarized by stating that if at least two separate and identifiable target features lie on the projection of the axis of rotation onto the imaging surface and do not move off this projection, then the reciprocal of the relative rate of dilation of the angle subtended by these features at the observation point (i.e., local tau along the line joining the features which is the same thing as local tau along the projection of the axis of rotation) is equal to the time-to-contact of the target with the observation point when the angle is small. If this condition is not met, local tau, however it is measured, cannot give the time-to-contact information described.

How might T_c with the observation point be obtained when two feature points lying on the projection of the axis of rotation cannot be located? Suppose an irreg-

ularly shaped object is rotating but is not changing its distance from the observer. The image will be continuously changing in size (and shape) but the *average* image size over complete revolutions of the object will be constant. If this object is now considered to approach the observer along a miss path whilst rotating and the *average* area of the image (or average solid angle subtended at the observation point by the object) over the i th complete revolution is \bar{A}_i , then the rate of change of the area can be approximated by

$$V_i^* = \frac{\bar{A}_{(i+1)} - \bar{A}_i}{\Delta t}, \quad (5.3)$$

where Δt is the time taken to make a measurement of average area (ie., the period of the object's revolution). The quantity τ_L might then be roughly approximated¹ by $2V_i^*/\bar{A}_i$. This quantity is by no means easy to come by, however. How, for example, does one ensure that the area has been averaged over a complete revolution? Information about how the object is rotating is required if appropriate estimates are to be made. Fortunately, this situation is not so complicated for a binocular observer for whom alternative information exists that does not require the extraction of τ_L (chapter 4) and it may be that monocular human observers find it very difficult to deal with the case that has just been considered (though there seems to be no direct empirical evidence on this matter). It is clear that the use of local tau is not as straightforward as is often implied (e.g., Solomon et al., 1984; Turvey & Carello, 1986).

¹ Lee & Young (1985) also remarked that some kind of temporal averaging was required in the extraction of time-to-contact information in certain difficult cases such as the ones discussed in this section.

§5.3. PROBLEMS IN EXTRACTING GLOBAL TAU

The perceptual measurement of global tau requires the location of the retinal coordinates of the focus of radial outflow of the optic velocity field. Location of these coordinates from an image velocity field which contains a rotational component due to eye movements has been the subject of much research (e.g., Lawton, Rieger & Steenstrup, 1987; Longuet-Higgins & Prazdny, 1980; Prazdny, 1981). It has been shown theoretically that a wide angle velocity field is important for accurate extraction of the FRO (Koenderink & Van Doorn, 1987) and recent empirical work indicates that people perform better at timing judgements when a wide angled field is available (Cavallo & Laurent, 1986; Groeger & Brown, 1988; see also chapter 6). Since the coordinates of the FRO coincide with the direction of heading the human error in locating the FRO is likely to be closely related to the error in detecting the direction of motion from optical flow patterns. Early work on the perception of direction of heading from optical flow (see review by Cutting, 1986) indicated that human observers were surprisingly poor at estimating their direction of translation from an optic velocity pattern, errors of several degrees of visual angle were reported in some studies (Cutting, 1986). More recent and carefully designed psychophysical studies conducted by Warren and colleagues (Warren & Hannon, 1988, 1990; Warren, Morris & Kalish, 1988) have shown that the direction of heading can, in fact, be located with a high degree of accuracy, estimated thresholds being in the region of 1° of visual angle or smaller. In this section the implications of an error of 1° of visual angle in location of the FRO for the accuracy of time-to-contact estimates provided by global tau will be considered.

In order to determine errors in time-to-contact estimated several quite strong assumptions will be made. These are as follows:

- (1) The error in estimating the direction of translation from an optical velocity pattern is directly proportional to the error in locating the FRO (for simplicity the constant of proportionality will be assumed equal to one).
- (2) Only the error in locating the FRO will be considered. This entails the assumption that this error does not interact with other errors involved in the computation of global tau.
- (3) The imaging geometry shown in figure 5.5 will be assumed.

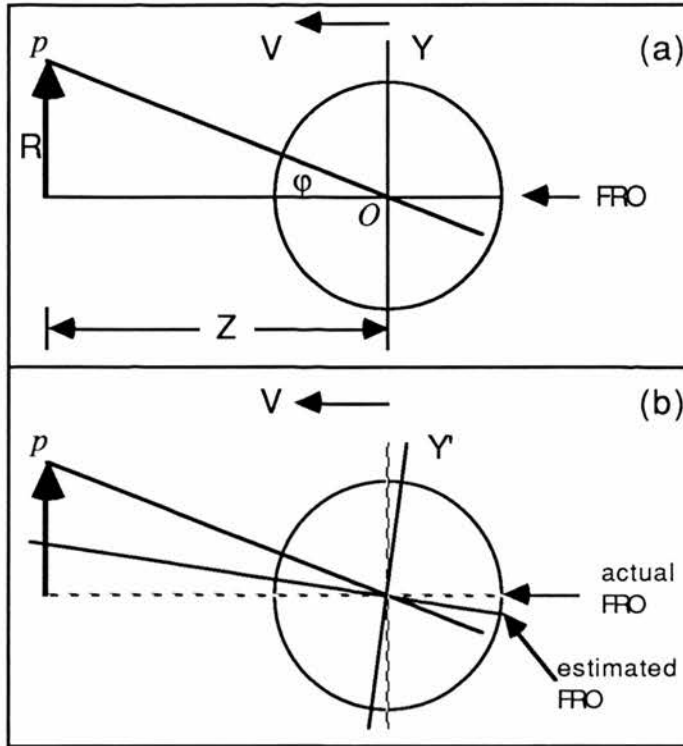


Figure 5.5. a) An imaging system with a spherical projection surface moves through a fixed environment with speed V . The induced image flow has a focus of radial outflow (FRO) at the point indicated. The Y axis is a "slice" through the plane perpendicular to the direction of motion through the observation point. b) There is an error in estimating the position of the FRO and the direction of travel from the image flow. The plane perpendicular to the estimated direction of travel is marked Y' .

Let (ϕ_0, θ_0) be the coordinates of the FRO on the unit sphere in some fixed coordinate system for labelling points on the sphere. As explained in chapter 4 (section 4.4, see equation 4.13) the following expression may be obtained for the time to contact with the Y plane, T_Y which equals Z/V_Z :

$$T_Y = \frac{\tan \phi - \tan \phi_0}{\dot{\phi} \sec^2 \phi} = \frac{\tan \theta - \tan \theta_0}{\dot{\theta} \sec^2 \theta}. \quad (5.4)$$

As noted in chapter 4 the right hand sides of equation 5.4 correspond to what was called global tau in section 4.4. Only the expression involving ϕ will be considered,

though that involving θ could clearly be used instead. Note that ϕ and $\dot{\phi}$ can be measured in our fixed image coordinate system but ϕ_0 must be computed from the measured image velocity field and it is errors in this computation that we are interested in.

For simplicity, the origin of the imaging surface coordinate system in figure 5.5a coincides with the FRO, i.e., $\phi_0 = 0$ (the essentials of the argument remain unchanged if this assumption is not made). In addition, suppose that the angular velocity $\dot{\phi}$ can be measured in these coordinates independently of locating the FRO (measurement of velocities is assumed to precede location of the FRO which is computed from the resulting velocity field). Errors in measuring ϕ and $\dot{\phi}$ will be ignored and assumed not to interact with the error in computing ϕ_0 which will be denoted ϵ_0 .

An error in locating the FRO will mean that time-to-contact will be estimated not with the plane Y in figure 5.5a but with plane Y' shown in figure 5.5b. If this time-to-contact is denoted by T'_Y then

$$T'_Y = T_Y \pm \frac{R \tan \epsilon_0}{V_Z}. \quad (5.5)$$

With $\phi_0 = 0$ from equation 5.5 we may write

$$\hat{T}'_Y = \frac{\tan \phi \pm \tan \epsilon_0}{\dot{\phi} \sec^2 \phi}. \quad (5.6)$$

where \hat{T}'_Y is the estimate of T'_Y provided by the visual system. Since $T_Y = \tan \phi / \dot{\phi} \sec^2 \phi$ and $\dot{\phi} \sec^2 \phi = -RV/Z^2$, equation 5.6 may be written

$$\hat{T}'_Y = T_Y \mp \frac{Z^2 \tan \epsilon_0}{RV_Z}. \quad (5.7)$$

Thus if the error ϵ_T in estimating T'_Y is equal to $T'_Y - \hat{T}'_Y$ then from equations 5.6 and 5.7 it is given by

$$\epsilon_T = \mp \frac{R \tan \epsilon_0}{V_Z} \left(1 + \frac{Z^2}{R^2} \right). \quad (5.8)$$

Consider the following example to illustrate what values for the error in time-to-contact are obtained by assuming errors in locating the FRO to be either 1° or 3° of visual angle based on expression for this error given by equation 5.8. Let an observer's task be to reach out and intercept a target object when it reaches a plane through the observation point perpendicular to the observer's estimated direction of motion

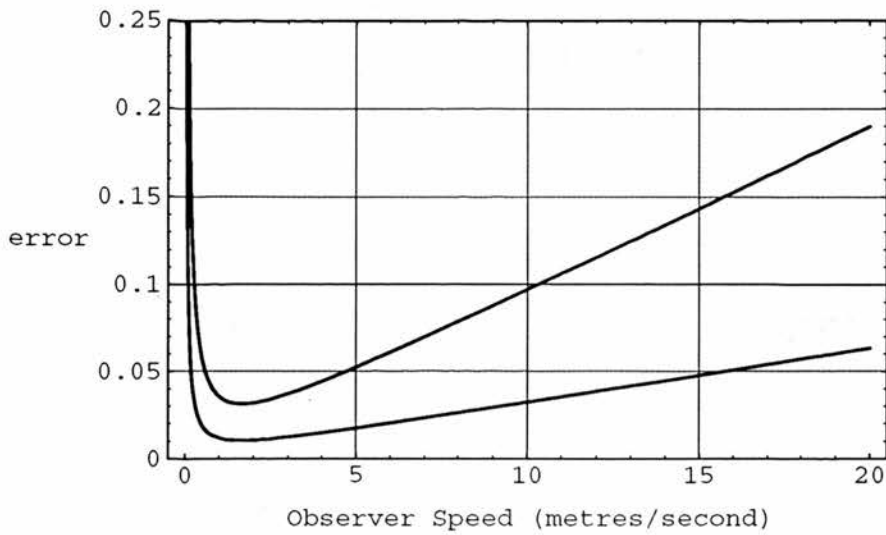


Figure 5.6. Timing errors (in seconds) at 300 milliseconds before contact as a function of observer speed due to FRO location errors associated with the use of global tau, calculated according to equation 5.8. Location error is 1° for the lower curve and 3° for the upper curve.

(the plane Y' in figure 5.5b). Assume that the time-to-contact information for this task is obtained 300 milliseconds before contact (Whiting and his colleagues showed that information obtained about 300ms before contact was critical in the execution of one handed catches but information obtained thereafter was less important, Sharp & Whiting, 1974, 1975; Whiting & Sharp, 1974). In addition, assume that the target object passes through the plane 50cms from the observer's eye. The error in the estimated time-to-contact is plotted against observer speed in figure 5.6 for FRO location errors of 1° and 3° . The graph shows that errors in time-to-contact estimations are significantly larger when the FRO location error is 3° than when it is 1° . In the former case timing errors are never lower than about 30 milliseconds which is outside the tolerance for successful catching as estimated by Alderson, Sully and Sully (1974). In the latter case, however, timing errors do not become prohibitively large except at very very slow speeds and relatively high speeds at which no one is likely to attempt the task described. In a range of speeds covering normal locomotor speeds the errors are in a tolerable range. For example, at an all out sprint, about 9 metres per second, the error is only 28 milliseconds or so and at a walking pace of 1.5 metres per second, the error falls to about 10 milliseconds, i.e., within the

tolerance for successful catching estimated by Alderson et al. It should be noted, however, that the error in locating the FRO is also going to be a function of observer speed. It is computationally more difficult to extract the FRO at very slow and very high speeds (cf Lawton et al., 1987; Koenderink & van Doorn, 1987; Warren, et al., 1988). Speculatively, it seems probable that error in locating the FRO will be a U-shaped function of observer speed, with the error largest at very slow and very fast speeds. This was not taken into account in figure 5.6 since no quantitative empirical data pertaining to this matter have been published. It may be tentatively suggested, however, that the error in estimating time-to-contact induced by error in locating the FRO is not great enough to impair successful interception of targets at speeds within the bounds for normal legged locomotion.

§5.4. MULTIMODAL AND INTERMODAL INFORMATION

Multimodal Information. It will be recalled from the discussion presented in chapter 1 that information was construed as being pattern or at least as being carried by pattern. Information was therefore described as independent of the modality of its transmission or intrinsically amodal. The information described in the last chapter was visual information in the sense that it was defined purely in terms of variables obtainable directly from the optic projection—observables of the optic projection. Denoting such visual observables by π_v , it will be observed that the expressions that were derived for time-to-contacts in chapter 4 were of the following general type:

$$E = f(\pi_v^{(1)}, \pi_v^{(2)}, \dots, \pi_v^{(n)}), \quad (5.11)$$

was critical in the execution of one handed catches but information obtained thereafter was less important (Sharp & Whiting, 1974, 1975; Whiting & Sharp, 1974).

where E represents the value of a physical (as opposed to perceptual) quantity. Written in this way with all the perceptual observables being visual, equation 5.11 represents purely visual information about E . If all the observables were auditory we would have purely auditory information. However, there is no reason why we should restrict our writing of equation 5.11 to purely unimodal expressions. We could, for instance, have an expression where some of the perceptual observables were obtained from one perceptual system and some from another, indeed we could have an expression in which all the perceptual systems contributed. When the perceptual variables π come from two or more different perceptual systems it seems appropriate to say that a relation of the form of equation 5.11 represents *multimodal* perceptual information.

At least two situations may arise. Firstly, we may have a multimodal expression in which the observables provided by the different contributing perceptual systems are only available to these perceptual systems and not to any of the others. Secondly, we may have a multimodal expression in which certain observables could be provided by several different perceptual systems. In the latter case the perceiver must either select which perceptual system is to be used as the source of a particular observable or use an appropriate estimate of the observable based on the measurements provided by the different perceptual systems to which it is available. This will be considered further below when we examine how this definition of multimodal information fits in with existing notions. First, however, we discuss how the ideas discussed above relate to the problems of perceptual timing information.

Multimodal and Intermodal Timing Information. Adopting the above notion of multimodal information, the timing information described by equations 4.21 and 4.23 represents potentially multimodal information. Consider the variable u (the rate of change of direction of the moving object measured relative to the 'eye') which appears in equations 4.21 and 4.23. A person wishing to catch an object will typically turn their head and generate pursuit eye-movements which act to keep the object's image on the foveal region of the retinae. This being the case, the rate of change of direction of the object is equal to the rate of rotation of the eyes relative to the environment if tracking is perfect. In illuminated environments such head-eye movements generate retinal flows which specify the rotation of the eyes in the environment—flows with a purely propriospecific significance (Koenderink

& van Doorn, 1981; Longuet-Higgins & Prazdny, 1980). In the dark no such flows are generated, of course. Nevertheless, if the head were to be fixed and only the eyes free to move then the rate of rotation of the eyes in the head as they follow the moving object is equal to the object's rate of change of direction. In this case the object's rate of change of direction is potentially available from extravisual sources since the rate of rotation of the eye in the head could in principle be obtained using proprioceptive/efference copy information. If the head is free to move, then the object's rate of change of direction is given jointly by the rate of rotation of the head on the shoulders and the rate of rotation of the eyes in the head⁵. Assuming that the object's rate of change of direction can be supplied by perceptual systems other than vision then the timing information described by equations 4.1 and 4.4 is potentially *multimodal* timing information.

Consider now the angular variable ψ which appears in equation 4.23. This quantity is the angle between the instantaneous direction of the moving object and the direction of the interception point measured at the observation point. To define ψ the *direction* of the point p is required. It would appear, therefore, that the direction of the point p needs to be visually available. It would be available if p were taken to be in the direction of some visible feature on a surface in the fixed environment; this feature then labels the direction of the interception point p . If such an environmental label for the interception point is required then the use of the information described by equation 4.23 would require that the observer have a structured visual environment. Under such conditions an observer wishing to intercept the object is able to do so along any direction from the point of observation through which the object will pass provided that some visible environmental feature lies in that direction. If it is supposed that the direction of the interception point is available from some extravisual source then the angle ψ is, at least in principle, available to an observer since both directions needed to define ψ are perceptually specified: the direction of the ball is visually specified, the direction of the intercept point is specified extravisually. In this case, therefore, it requires two sources of information to define ψ . If both sources are perceptual then ψ is defined *across* perceptual systems as it were. In this case it seems appropriate to say that ψ is *intermodally* specified—it is defined as a relationship *between* perceptual systems.

⁵ Note that there will be some translation of the observation point when the head rotates and indeed when the eye rotates. It will be ignored here for the sake of illustration, but would have to be taken into account if u is to be accurately determined.

It will be observed that any intermodally specified quantity can be expressed as a purely multimodal relationship of the form given in equation 5.11. In the example given above the angle ψ is derived from a direction measured in a proprioceptive frame of reference and a direction measured in a visual frame of reference. Knowing the relationship between these two frames allows ψ to be computed. However, when we have a variable such as an angle like ψ which is unitary in the sense that it could be computed directly by a single perceptual system, the term *intermodal* seems most appropriate.

There have been at least two ways, in addition to that described above, that perception might be considered multimodal. One of these concerns the placing of quantities defined in a sensory frame of reference, e.g., defined in a retinal frame of reference in the case of vision, into a frame of reference that is independent of the sense organs themselves. Such processes are required if we are, for example, to see something as having a fixed position in space rather than as moving about every time we move our eyes. Establishing a sensory independent frame of reference for perception may require the cooperation of several perceptual systems. This problem is discussed in detail in the next section.

Other Notions of Intermodal and Multimodal Perception. Perception is sometimes considered to be multimodal when the *same* information is potentially available to more than one perceptual system. One of the most extensively analyzed examples is the light-gravity orientation of fish. It was studied by von Holst and his colleagues (original references in German, see Schöne, 1984) as well as by others (Braemer & Braemer, 1958; Stange, 1972). Observation of certain types of fish in an aquarium near a bright window reveals that they often assume a tilted position with their backs leaning toward the light. Von Holst and his colleagues studied this phenomena using a method which involved changing the intensity of the gravito-inertial force vector by rotating the aquarium about a remote axis using a specially designed "centrifuge". They showed that the fish orient their normal position simultaneously to light and gravity with the position corresponding to the vector resultant of gravity and light (both having a direction and an intensity). The fish were found to weight gravity and light differently under different conditions. For example, von Holst found that when hydrostatic pressure is increased, fish weight gravity more strongly, whereas in rough water they weight light more strongly. Schöne (1984) suggests that the biological significance of these changes in the rela-

tive weighting of gravity and light is (1) that higher pressure tends to be correlated with deeper water where there is less light and the fish must consequently rely more on gravity for its orientation, and (2) the reliability of labyrinth function (the labyrinths signal the direction of the gravito-inertial force vector) is diminished in rough water when the fish is tossed about and must consequently rely more on vision for its orientation.

This example illustrates the types of problem that arise when different perceptual systems supply the information about the same quantity. As described, it may happen that one perceptual system supplies more reliable information than another. In such a situation the animal is faced with the problem of what relative importance to attach to each perceptual system and how this relative importance changes with context. The organism may thus be considered as being equipped with context sensitive weighting schemes which are based on the relative importance or reliability of different sources of information.

It should be pointed out that this kind of problem does not only arise when the same information is supplied by different perceptual systems. There may be different sources of information about the same quantity or property within a single perceptual system as exemplified by the famous shape-from-x problems in low-level vision where x might be shading, texture, stereo or motion (Aloimonos, 1988; Horn & Brooks, 1989). This is clearly related to classical cue theory which describes, for example, different cues to depth and so on. Cue theory was based on the notion of informationally impoverished stimuli which is nowadays considered an inappropriate characterisation of the perceptual input (chapter 1) but the term 'cue' is now usually used to mean something which acts as a source of information. Just as for the multimodal case, a single perceptual system with several different sources for the same information is faced with the problem of which source to use or with how to combine them to yield a reliable representation of the property that is being informed about. Problems of this kind are currently being investigated extensively by psychophysicists who examine how effectively different cues are used to perform visual discrimination tasks (as assessed by threshold measurements) and whether, in the presence of more than one type of cue, one dominates or whether they combine to give improved discrimination (i.e., lower thresholds than any one cue gives on its own).

The notion of multimodal information defined as information available to more

than one perceptual system is not the same as the notion that was introduced at the beginning of this section but, as indicated above, is complementary to it. The former notion, by considering the same information to be available to different perceptual systems, introduces the problem of the relative weighting of the different sources of information. The notion of multimodal information that was introduced at the beginning of this section, however, regards certain information as being defined jointly by two or more perceptual systems. A question that one might ask at this stage is why use the term 'multimodal information' for the notion introduced here when it is already in use? The reason is that the original use of the term refers to information that is simply *available* to more than one sensory system whereas the term is used here to refer to information that is actually *defined* in terms of variables obtained from more than one sensory system. It seems natural to call information that is defined multimodally, multimodal information and use a different term to refer to information that is available to more than one sensory system e.g., "multimodal cue".

Stoffregen and Riccio (1988) suggested a notion of "intermodal" information that corresponds to the concept of multimodal information introduced earlier. Although they did not formulate the concept for the general case it is clear that they had the same thing in mind. They based their formulation on the intuitions of the Gibsons who suggested the basic notion of multimodal information in the sense used here (see e.g., E. J. Gibson, 1969). Stoffregen and Riccio present the following equation relating three perceptually available variables, θ , $\ddot{\theta}$ and T_c (the latter representing torque and not time-to-contact in Stoffregen and Riccio's example):

$$\theta(t) = k_1 \ddot{\theta}(t) + k_2 T_c$$

where k_1 and k_2 are constants (Stoffregen & Riccio, 1988, equation 6). In relation to this equation Stoffregen and Riccio observe that "[b]ecause $\ddot{\theta}$ and T are available to different perceptual systems, the ... information ... is of an intermodal nature" (pp 11). Clearly Stoffregen and Riccio are talking about what we earlier called multimodal information in the specific context where their equation 6 is relevant (it actually describes information for postural orientation but this need not concern us here). These authors take this notion of multimodal perception as standing in direct opposition to the other notion of multimodal perception discussed above. The view taken here, however, is that the two notions are complementary which allows for the possibility that information which may under normal conditions be

obtained unimodally may sometimes be obtained multimodally (or intermodally) if the normal unimodal source becomes impoverished for some reason (see the above discussion of timing information and chapter 7).

§5.5. CONCLUSIONS

In this chapter the problems that are involved in actually extracting the optic variable τ were discussed. It was shown that problems arise in using local τ as information about time-to-contact of a target object with the observation point when the object is non-spherical and rotating relative to the observer. These problems can, in principle, be overcome by measuring local τ in different ways, either along a single spatial dimension in the retinal image or in terms of temporally averaged quantities. The kind of measurement that is appropriate depends upon the context and may require additional information (such as the axis of rotation of a moving object) suggesting that the extraction of time-to-contact information cannot be as simple and low-level a process as is sometimes implied. The visual system needs to determine what quantity needs to be measured in a particular situation and to act accordingly if it is to extract time-to-contact information effectively. Unfortunately, very little is known about how the human visual system deals with the cases discussed in section 5.2 of this chapter which complicate the extraction of local τ .

Global τ does not suffer from the same difficulties as local τ but it cannot be used as information about time-to-contact when the object of interest is moving relative to the environmental frame of reference which is, of course, a very serious limitation. It can only be used as time-to-contact information about fixed surfaces during locomotion of the observation point. In order to use global τ , therefore, the surfaces of interest have to be identified as being stationary with respect to the environment. When surfaces are stationary a moving visual system has the choice of using global or local τ .

The general time-to-contact information described in chapter 4 requires that other variables in addition to local τ be available. In this chapter the possible sources of these variables were discussed and it was shown that they need not be extracted visually since they are potentially available to other sensory systems. In this context, specific notions of multimodal and intermodal perceptual information were defined and compared with existing uses of these terms. It was argued that

the definitions given here correspond more closely to what the terms appear to mean intuitively than do other uses. In the next chapter existing empirical data about the extraction of time-to-contact information by the human visual system is considered and outstanding empirical questions are identified.

6

Empirical Investigations and the Distance÷Velocity Hypothesis

§6.1. INTRODUCTION

This chapter discusses recent experimental investigations of the perception of time-to-contact. These investigations have been largely motivated by a single theoretical question: is time-to-contact perceived through the optic variable tau or does it require computation from independently available distance and velocity? Interest in this question seems to derive from the characterisation of the tau account as an example of direct perception and the distance÷velocity account as an example of indirect perception. As should be apparent from the discussion of earlier chapters, however, this distinction is of little significance. The question is potentially of interest because it concerns the source of timing information actually used by the visual system, given that it could use tau-based information or obtain it from distance and velocity information through division.

In this chapter arguments are presented which lead to the following conclusions:

- Published experimental data cannot be interpreted as having any bearing on the issue of whether a tau-based or a distance÷velocity method is used despite the fact that it is routinely interpreted in this way.
- The distance÷velocity method makes an implicit assertion about the processes involved in extracting time-to-contact information, an assertion which should be justified before the distance velocity method could be taken seriously as an account of the perception of time-to-contact.

In the discussion which follows it will be assumed that people are able to correctly interpret local tau. Some of the examples given in this chapter involve ball games and since most balls are spherical problems interpreting local tau will not arise. When problems do arise it will be assumed that they can be solved.

§6.2. WHAT TIMING INFORMATION IS USED?

Two cases will be discussed in this section:

1. An object is on a collision course with an observer's eye and the observer interacts with the object a small distance in front of the eye.
2. An object is not on a direct collision course but rather passes by a short distance away from the observer who intends to interact with it whilst it is within reach.

In these two cases the relevant tau variable is local tau since the object is in motion relative to the fixed environment. Consider the first case: the optic variable tau specifies T_c of an approaching object with the observation point (the eye). However, if the object is, for example, a baseball heading directly towards the head of an outfielder the fielder will catch it at a short distance in front of him. Consider the situation represented schematically in figure 6.1.

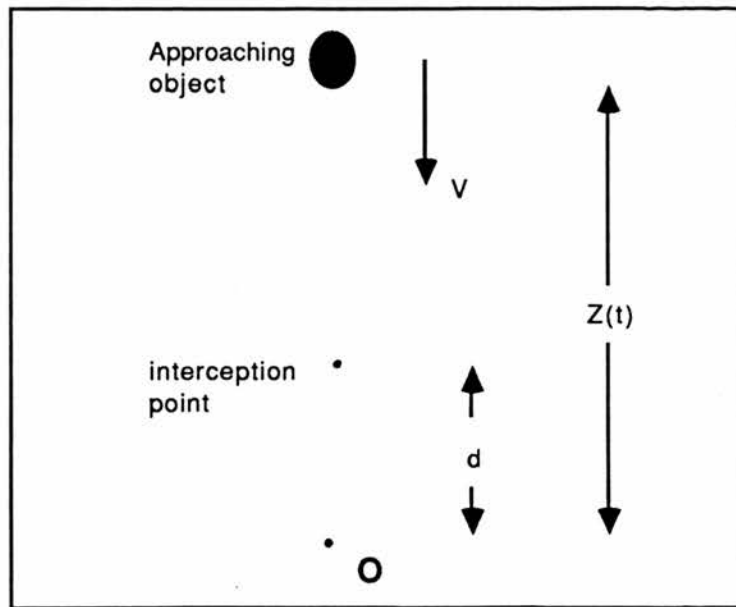


Figure 6.1. A moving object is approaching a point of observation O on a collision course with speed a constant V . It is instantaneously a distance $Z(t)$ from O. The interception point is a distance d in front of O.

Referring to the figure, the time-to-contact with the eye (Z/V , where V is the speed) at an instant of time is given by the value of τ_L at that time. Time-to-

contact with the interception point, T_i , is given by,

$$T_i = \frac{(Z - d)}{V}, \quad (6.1)$$

which can be written

$$T_i = \tau_L - \frac{d}{V}. \quad (6.2)$$

Thus, taking τ_L as an estimate of T_i introduces an error equal to d/V . Suppose that the catcher's arm is outstretched such that the distance, d , of the hand from the eye is 50cm. If the ball were travelling at $8ms^{-1}$ then the margin of error for a successful catch is about that estimated by Alderson et al. (1974), i.e., of the order of 15ms or so. Under these conditions the error induced by taking τ_L as an estimate of time-to-contact is 62.5ms. Skilled catching with the arm outstretched would therefore be impossible, it would be fumbling and incompetent at best. Skilled timing of the catch would only be possible at distances of less than about 12cm in front of the eye. This does not accord well with everyday experience of catching skills and the reader is encouraged to try catching a soft ball falling towards his or her head with outstretched arm. It seems as though one needs not only τ_L but information about distance, d , to the interception point and the speed V of the approaching object if timing is to be precise.

The situation just described has some bearing on the theory of braking control proposed by Lee (1976). Lee suggests that the rate of change of τ , i.e., its temporal derivative $\dot{\tau}$, can be used to inform about whether a car driver's deceleration is adequate to stop in time to avoid collision with an upcoming obstacle¹. Lee argues as follows: applying Newton's equations he deduces that the driver's deceleration (D) is adequate if and only if the distance it will take the car to stop with that deceleration is less than or equal to the current distance (Z) from the upcoming obstacle, i.e., if and only if

$$\frac{V^2}{2D} \leq Z, \quad (6.3)$$

where V is the car's instantaneous speed. From this Lee deduces that deceleration is adequate if and only if

$$\dot{\tau} \geq -0.5, \quad (6.4)$$

and concludes that, "a safe braking strategy would therefore consist of the driver adjusting his braking so that $d\tau(t)/dt$ remained a safe value" (Lee, 1980a; pp. 294).

¹ Note that global tau is the relevant tau-variable in this case.

This conclusion that the strategy is safe is unwarranted, however, since no account is taken of the extent of the car body in front of the driver. Suppose the car extends a distance d in front of the driver, taking this into account requires that Z in equation 6.3 be replaced with $Z - d$, the distance of the obstacle from the front of the car. Thus, the following equation is obtained

$$[Z(t) - d]D/V^2(t) \geq 0.5. \quad (6.5)$$

By differentiating the relation $\tau(t) = Z(t)/V(t)$ one gets,

$$Z(t)D/V^2(t) = 1 + \dot{\tau}(t). \quad (6.6)$$

From equations 6.6 and 6.5 the following expression is obtained,

$$\dot{\tau} - \left(\frac{\dot{\tau} + 1}{\tau} \right) \frac{d}{V} \geq -0.5, \quad (6.7)$$

As Lee has pointed out (personal communication) the quantity d/V in equation 6.7 is the time-to-contact with the plane of the driver's eye of the "piece" of road instantaneously next to the front of the car. This T_c is specified by the value of *global* tau for a texture element lying on this piece of road. Thus, denoting the tau value for the obstacle by τ_o and the tau value for the texture element on the road by τ_r , equation 6.5 may be written entirely in terms of optical variables as follows:

$$\dot{\tau}_o - \left(\frac{\dot{\tau}_o + 1}{\tau_o} \right) \tau_r \geq -0.5. \quad (6.8)$$

Thus, in taking into account the extent of the car in front of the driver it is not enough to use simply the quantity $\dot{\tau}$ as Lee originally suggested. Adopting Lee's braking strategy entails making errors in braking which are likely to lead to crashes since the driver's tendency will be to brake too late. It cannot be concluded, therefore, that Lee's strategy is a *safe* braking strategy.

What happens if one attempts to use tau as an approximation of T_c when the object is not on a collision course with the eye? The situation represented in figure 6.2 will be considered. In this case the relation between τ_L (under spherical projection) and the illustrated environmental quantities, derived in von Hofsten & Lee (1985), is

$$\tau_L = T_N + S^2/ZV, \quad (6.9)$$

where T_N is the time-to-contact with the point of nearest approach. Assume that the observer wishes to intercept the object at the point of nearest approach (N).

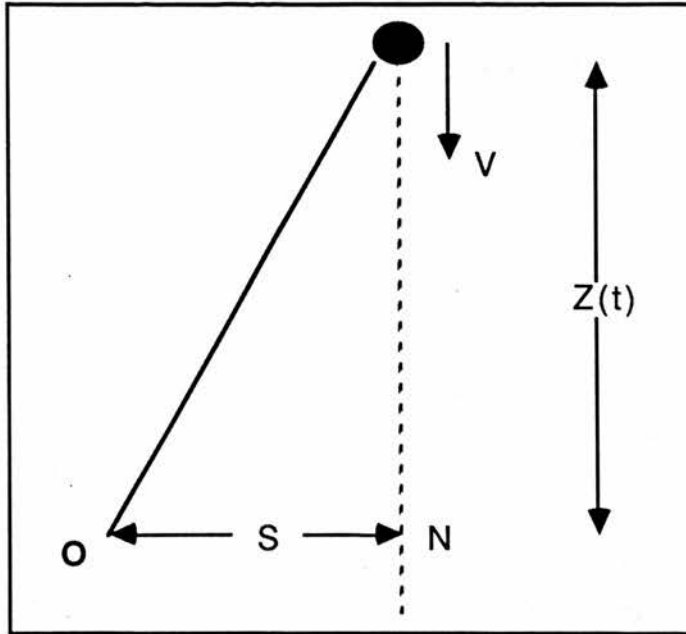


Figure 6.2. A ball moving in the plane of the page along the path shown as a dotted line is approaching the observation point O with constant speed V . It is instantaneously a distance $Z(t)$ from N , the point of nearest approach to O . Time-to-nearest approach is equal to $Z(t)/V$.

This seems a sensible point to consider since if the object is intercepted in front of N (nearer the object in figure 6.2) the error involved in taking τ_L as an approximation to T_c will be slightly less than that at N whereas if the object is to be intercepted behind N the error will be slightly greater. The error introduced by using τ_L as an approximation to time-to-contact in this situation is of magnitude S^2/ZV . Is this error acceptable? If S is small or V large it may be satisfactory. Consider a slip fielder in the game of cricket. Suppose he is faced with the problem of catching a ball travelling at $20ms^{-1}$ which will pass to his right at a catchable distance, say 50cms. Let us suppose further, based on the findings of Whiting and his colleagues (e.g., Sharp & Whiting, 1974, 1975), that information obtained 300ms before contact is adequate for reasonably effective catching. The error in taking τ_L as an estimate of T_N may be calculated² to be about 2ms. Thus the tau approximation supplies very precise information indeed under these conditions. It is unlikely that a biological vision system could calculate τ_L accurately enough to

² Since $Z/V = 0.3$ seconds and $V = 20ms^{-1}$, $Z = 0.3 \times 20$. The error S^2/ZV is equal to $0.5^2/0.3 \times 20 \times 20$ which comes out to 2.08milliseconds.

exploit this precision but it may not need to — a temporal accuracy of between 5 and 10ms is probably what is required to catch effectively in the cricket slips (Lee & Young, 1985).

When the ball is moving at the more moderate speed of $6ms^{-1}$ the error increases to 23ms ($S=50cms$), but this may still be just about good enough to achieve reasonable catching success (Alderson et al., 1974; Sharp & Whiting, 1974). It is at slower speeds that taking tau as an approximation begins to look really suspect. When the object to be caught moves at a speed of $4ms^{-1}$ the error ($S=50cms$) increases to 52ms and at $1ms^{-1}$ it is 800ms. Indeed, using ball speeds of less than $1ms^{-1}$ von Hofsten (1983) found that the catching skills of very young infants were far more accurate than a tau approximation strategy could account for.

It may be concluded that at very fast object speeds a tau approximation strategy is effective but it is less so for slow and moderate ball speeds. At these speeds more precise information is required such as that supplied by time-to-nearest approach information (Lee & Young, 1985) or the more general timing information described in chapter 4, both of which require the detection of perceptual variables other than τ_L .

We have attempted to show here that to account for the skilled timing of interceptive action a strategy based solely on tau is inadequate. To accurately time interactions, even when the object to be intercepted or avoided is on a collision course with the eye, requires information in addition to that provided by τ_L . It is largely because of this fact that existing studies shed no light on the question of whether a tau-based or a distance÷velocity strategy is employed. Many published studies attempt to show that information in addition to τ is required in T_c estimation and use this as evidence for a distance÷velocity strategy (e.g., Cavallo & Laurent, 1988; Cavallo et al., 1986; Groeger & Brown, 1988), but it clearly does not constitute such evidence.

§6.3. THE DISTANCE÷VELOCITY ACCOUNT OF T_c PERCEPTION

It is possible that an observer could obtain T_c information not from a relation which specifies it directly without need for further computation such as those discussed in the previous section, but from perceived distance and velocity information. If the observer can obtain information about the distance of the moving target from the interception point and its velocity then (assuming velocity remains constant) T_c of the target with the interception point could be obtained by dividing this distance by this velocity³. Although this method for obtaining T_c information is frequently cited as an alternative to a tau based account, no fully worked out distance÷velocity scheme is to be found in the literature. What needs to be specified is precisely what perceptual information about real world distances and velocities might be used to compute T_c .

There is evidence that T_c can be estimated when no absolute distance and velocity information is available (Schiff & Detwiler, 1979; Todd, 1981), which may be taken as evidence for the use of a tau-based strategy. This does not establish that when target distance and velocity information is available it is not used to compute T_c , neither does it exclude the possibility recently suggested by Schiff and Oldak (1990) that in some circumstances tau-based strategy is used and in others distance÷velocity strategy is used, specifically situations where tau is below threshold. However, as was shown in chapters 4 and 5 there is time-to-contact information available in non-collision approaches which does not require local tau to be detected.

The possibility remains, of course, that distance and velocity are used to obtain T_c information when they are available. A variety of possibilities exist as to the source of such information. A monocular observer could make use of optical information about distance and velocity (described below) or perhaps ocular accommodation cues. A binocular observer could also use binocular disparity information (appropriately scaled) or binocular convergence angle (again appropriately scaled) to obtain information about the objective distance of the target. Several researchers have interpreted experimental results as indicating the use of a distance÷velocity strategy in certain conditions (e.g., Cavallo & Laurent, 1988; Groeger & Brown,

³ Discussion in this section will continue to assume constant velocity for simplicity (the argument that will be developed remains essentially unchanged if acceleration is taken into account, as should be clear from appendix A).

1986; McLeod & Ross, 1983). The data, however, do not support such an interpretation, a conclusion which will be discussed fully in the next section. Nevertheless, does such an account provide a plausible and testable alternative to a tau-based account?

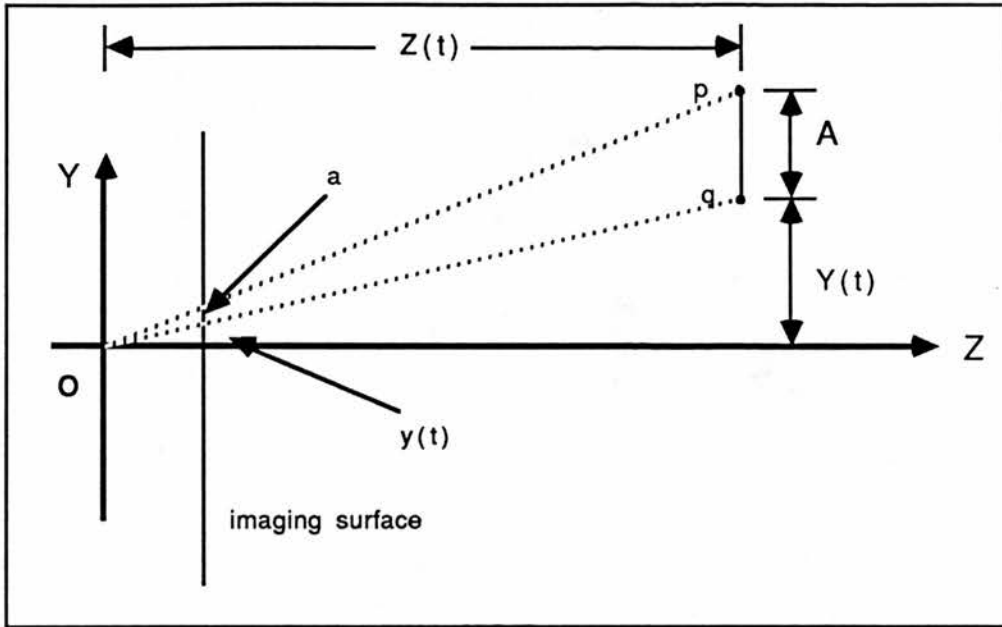


Figure 6.3. Todd's (1981) model of the imaging situation; it can be considered a 'slice' through the imaging system of Longuet-Higgins and Prazdny (figure 4.4). This is a reasonable idealization since in the real world objects tend to move in a plane. The imaging system is fixed relative to the environment and the line segment length A joins two points p and q of an object moving through the environment in the YZ plane.

It will be instructive to consider Todd's (1981) analysis of the situation illustrated in figure 6.3. The following relations can be derived quite easily (see Todd, 1981):

$$Z/A = 1/a, \quad (6.10)$$

$$\dot{Z}/A = -\dot{a}/a^2, \quad (6.11)$$

$$Y/A = y/a, \quad (6.12)$$

$$\dot{Y}/A = \dot{y}/a - y\dot{a}/a^2. \quad (6.13)$$

These equations show that visual information about the (instantaneous) velocity and position of the moving object measured relative to the Z - Y coordinate system

is available since these quantities scaled in terms of the object size A are given in terms of velocity field variables. If a perceiver has prior knowledge about the object size then he or she can obtain useful information about the position (Z, Y coordinates) of the object and its velocity. Given equations (6.10) to (6.13) it is a simple matter to show that T_c information is available. Dividing equation (6.10) by (6.11) one obtains an expression for $-Z/\dot{Z}$ —the T_c with the Y -axis:

$$-Z/\dot{Z} = a/\dot{a}. \quad (6.14)$$

A similar operation on equations (6.10) and (6.11) gives the time-to-contact with the Z -axis. The right hand side of equation (6.14) is local tau (τ_L). The expression for target velocity can then be written $V_Z/A = 1/a\tau_L$ indicating that τ_L could play a role in the extraction of quantities other than T_c . It has thus been shown that if a monocular observer has prior knowledge about the object size A then he or she can obtain information about the object's distance away and its velocity.

The derivation of an expression for T_c (equation 6.14) can be read in two ways: either as a purely formal argument showing that T_c information is visually available or as an account of how the visual system might actually compute T_c . Read in the latter sense we get the following account for T_c perception by a monocular observer: the observer's visual system computes distance and velocity according to the two expressions given above (it 'knows' about the object's size, A) and then divides these two perceived quantities to obtain T_c . The alternative account says that the observer's visual system simply measures τ_L . It does not need to obtain it via the roundabout method of measuring $1/a$ and $1/a\tau_L$ scaling these two quantities with respect to prior information about the object's size and then dividing the results of these two processes to get exactly the same result as would be obtained by measuring τ_L . (In fact, this roundabout method could be viewed as a means of computing τ_L .)

Such a distance÷velocity account clearly begs the following question: why should the visual system ever take the roundabout distance÷velocity route (which is probably more prone to error) to T_c information when exactly the same information could be extracted in a much simpler way? It is difficult to see what the rationale for such a route would be. This problem is not only a feature of the monocular formulation of the distance÷velocity strategy presented above, a similar problem arises for any such strategy. The reason for this is that to obtain information about the distances and velocities required to compute T_c , the observer

must scale the *relative* depth and velocity information present in the stimulus. For example, in the case of an object some distance away but directly in front of the observer, binocular disparities and convergence angles must be scaled by the interocular separation if the distance away is to be obtained from either of these sources. Thus any distance÷velocity strategy first scales stimulus information and then undoes this scaling to get T_c information. Whilst a perceiver might possibly be driven to use a distance+velocity strategy under unusual circumstances, without some guiding rationale it is difficult to argue that such a strategy is used in normal circumstances where it is an uneconomic method for computing T_c . Nonetheless, as noted above, certain empirical results have been interpreted as evidence for the use of a distance÷velocity strategy by human observers. In the next section the interpretation of empirical results is considered.

§6.4. EMPIRICAL STUDIES

It is frequently argued that if it could be shown empirically that human observers used more than tau alone to judge T_c this would go against a tau-based account of the perception of T_c and tend to support instead some kind of “computational” account presumably based on distance and velocity (e.g., Cavallo & Laurent, 1988; Groeger & Brown, 1988; McLeod & Ross, 1983; McLeod et al., 1985). The discussion presented above demonstrates that evidence for factors other than tau in T_c judgements is precisely what is expected from a tau-based account extended to deal with the real world requirements of interceptive timing. This is also true if it were found that binocular vision improved accuracy of time-to-contact estimation. Several authors have interpreted improved timing performance with binocular vision over that obtained with monocular vision as evidence in favour of the distance÷velocity strategy (Cavallo & Laurent, 1988; McLeod et al., 1985). It should be clear from the discussion of the last two chapters that better performance with binocular vision does not constitute such evidence. As shown in chapter 4, there is binocular information specifying T_c which may be superior to monocular information under certain circumstances (see chapter 5). In addition, Jones and Lee (1981) reported tasks in which “two eyes were better than one” where, apparently, binocular depth information was unnecessary. It may be that two eyes enable better extraction of information which is defined monocularly — two eyes give you two goes at the same thing, as it were (see Jones & Lee, 1981, for

relevant discussion). In this section, studies by Cavallo and Laurent (1988) and by Schiff and Detwiler (1979) are discussed, both of which have been interpreted as bearing on the issue of whether a distance÷velocity or tau-based strategy is used to obtain T_c information. It will be argued that the results cannot be interpreted as distinguishing the two strategies.

Cavallo and Laurent (1988) reported data collected from subjects who were driven at a target in an automobile. Shortly before contact with the target, subject's vision was occluded and they were required to press a button at the moment they expected the car to collide with the target. A major finding was that the extent of the subject's visual field had a significant effect on T_c estimation for inexperienced not for experienced drivers. The larger the field of view the better the estimates made by the inexperienced drivers. It was proposed that if tau information alone were being used then one would expect the results to be independent of the size of the visual field. This is perhaps true for local tau but false for global tau, because, as shown theoretically by Koenderink and van Doorn (1987), the accuracy with which the focus of radial outflow can be localised depends on the field of view: in general it will be better localised the larger the field of view.

However, the authors observe that even if a case could be made for an enhancing effect of a wide visual field on the pick-up of tau one would expect this to be independent of driving experience. They conclude that the visual field effect is attributable to the fact that beginners assessed speed as a separate parameter and may thus be interpreted as using a distance÷velocity strategy, whereas the performance of the experienced drivers is consistent with a tau-based strategy. This explanation of the visual field effect raises the question of why the novice drivers should use a completely different source of information from the experienced drivers? The analysis of the tau account presented above suggests two alternative explanations which do not suffer from this difficulty. First, it is possible that driving experience makes the extraction of global tau more robust against decreases in the size of the visual field. Second, returning to the analysis presented earlier with reference to figure 6.1, it is clear that in order to judge accurately when the front of the car will hit an obstacle the velocity of the car and the extent of the car in front of the driver need to be taken into account (see equation 6.7). It is then possible to argue that when the field of view is restricted and vision of the bonnet is occluded the driver no longer has access to information about the length of

the bonnet. It would not be surprising if the experienced drivers had learnt about hood lengths through extensive experience and could take it into account without having to actually see the hood itself whereas novice drivers needed to see the hood in order to take its length into account effectively. It might be expected, therefore, that the difference between novice and expert drivers would decrease as the speed of approach to the target increased since the error induced by failing to take the hood length into account is negligible at high speeds. Thus the differential effects of field of view would only be evident at rather slow speeds given the precision of T_c estimation shown by subjects in this kind of task (note that this is not a prediction of the first explanation). Cavallo and Laurent do not provide data that would allow this prediction to be tested.

Schiff and Detwiler (1979) presented subjects with short animated film sequences showing a black square directly approaching the camera position. The background was divided into two regions (terrain and sky) by a thin black "horizon" line. Either both terrain or sky were plain white or one was white and the other covered in a grid of squares. Schiff and Detwiler claimed that the grid terrain provided subject's with enhanced distance and "distance change" (velocity) information. However, there was no *absolute* distance or velocity information in Schiff & Detwiler's stimuli and such information is required for the distance÷velocity strategy. The reason for the absence of absolute distance and velocity information is clear — there was nothing in terms of which the *relative* distance and velocity present in the stimuli could be scaled. Since the subjects were watching a two dimensional film there was no binocular depth information and they also had no means by which they could scale the scene presented in the films — it could be a small scene near by or a large one a long way off. Thus, the results of this study can shed no light on whether tau or distance÷velocity is used to obtain T_c information. The data do indicate that subjects are able to use image expansion (in the form of τ) when both distance and velocity information are unavailable.

§6.5. IDENTIFYING EMPIRICAL PROBLEMS

From the discussion presented so far it may be concluded that no existing empirical research sheds any light on the question of whether a distance divided by velocity strategy or a tau based strategy is used to obtain T_c information. Results which are routinely interpreted as favouring the distance÷velocity strategy are exactly what would be expected from a tau based strategy which is general enough to deal with the requirements of natural skilled timing. Indeed, as discussed in the section 3, it is difficult to formulate a version of the distance÷velocity strategy which does not make assumptions about the visual system which have no empirical or theoretical support.

However, still of potential interest is the question of whether, when tau is unavailable, other information could be used to estimate T_c . Both Schiff and Detwiler (1979) and Todd (1981) report results showing that human subjects can estimate T_c in the absence of any distance or velocity information. It would be interesting to investigate the converse: can subjects reliably estimate T_c when both distance and velocity information are available but tau is not (this would include tau defined in the most general sense as described in chapter 3 so as to exclude use of the binocular information). However, it is difficult to see how one could provide both distance and velocity information and at the same time avoid providing tau as well. As is clear from the discussion presented in the previous section, providing monocular information about distance and velocity entails providing tau. Indeed, the situation as far as visual information about T_c is concerned seems to be that providing sufficient information for a perceiver to be able to compute the relevant distances and velocities to obtain a desired T_c entails providing the corresponding tau based information about that T_c . In contrast to this, the reason that it has been possible to investigate whether T_c can be estimated when information about distance and velocity is absent and tau is present is precisely because providing tau does not entail providing distance and velocity. The thrust of the last section was simply that if visual information about both distance and velocity is available then so is the corresponding tau based information and it is difficult to justify the claim that a visual system should not be able to exploit this fact. Given the analysis presented here and in the last section, it seems as though the issue of what strategy is used to obtain T_c information cannot be formulated well enough to lead to fruitful empirical questions. In addition, it is worth remarking that much of the

motivation for investigating this issue has been to contrast a "Gibsonian method" (τ) with a "computational method" (distance divided by velocity) (e.g., Cavallo & Laurant, 1988; McLeod & Ross, 1983). In chapter 1 it was argued that there is really little meaningful distinction to be drawn between computational approaches and Gibsonian ideas and τ based methods for obtaining T_c information are no less "computational" than any of the work in computational vision (see chapters 4, 5 and Koenderink, 1985). Consequently, the issue of what strategy is used to obtain T_c loses much of its *raison d'être*.

What other empirical questions can be raised concerning the perception of time-to-contact? The analysis presented in this and the preceding two chapters poses a number of interesting questions which will now be detailed.

(1) When an object is approaching along a miss path does a person make use of the more general T_c information described in chapter 4 to time an interaction with it? In particular we can enquire as to whether the person uses u (the rate of change of direction measured at the point of observation) and ψ (the angle between the instantaneous position of the moving object and the interception point subtended at the observation point).

(2) Can timing information be obtained multimodally and intermodally as described in chapter 5?

(3) When an approaching object is irregularly shaped such that changes in retinal image size can result from changes in the direction of the object from the eye or from rotation of the object about some axis passing through it, is monocular estimation of T_c degraded relative to that for a spherical object? If estimation were significantly degraded this would indicate that the methods for discounting changes in image size not due to approach to the observation point described in the last chapter are not being used (or being used efficiently) by the visual system.

(4) Can a person use the binocular information about T_c described in chapter 4? This information is potentially valuable when the moving object is irregularly shaped and spinning about an axis perpendicular to the plane of its motion. In this case, discounting changes in image size not due to approach to the eye would require some kind of temporally extended procedure (see chapter 5). Using the binocular information would therefore be potentially faster and possibly less sen-

sitive to noise.

(5) When a moving object is on a collision course and interception is to be achieved some distance in front of the eye, is the distance and velocity information described in section 1 of this chapter used (in addition to tau) to time the interception? There is also the related question of whether bonnet length and velocity are used in braking as suggested by the results of Cavallo and Laurent (1988) or whether tau for the "piece" of road a bonnet's distance from the driver is used instead (see above). It is possible that when the road is occluded bonnet length and velocity can be used but when the road is visible the appropriate tau value is used instead.

(6) The strategy of assuming constant velocity of approach when timing interceptive actions has been tested only in the case of more or less direct approaches to the eye (Lee & Reddish, 1981; Lee et al., 1983; Todd, 1981). It has not been tested for miss path approaches where the more complex T_c information described in chapter 4 is relevant.

These six questions are the most salient of those raised by the analysis of the perception of time-to-contact presented in this thesis. In the next chapter experiments are described which sought to investigate aspects of questions 1 and 2 above. In the final chapter experimental methods for investigating some of the other questions listed above are suggested.

7

Experimental Investigations

§7.1. INTRODUCTION

In this chapter the results of several experiments are reported that attempt to discover what other perceptual variables apart from image expansion are involved in the perception of time-to-contact when a target is not directly approaching the observation point. The analysis presented in chapter 4 identified two variables as potentially important in accurate perception of time-to-contact during such approaches (assuming a constant velocity strategy). These variables were a) the instantaneous angle subtended at the observation point by the target and the interception point together and b) the rate of change of direction of the target. In chapter 5 the possibility was raised that these variables could be obtained multimodally or intermodally (i.e., in a manner that involved contributions from more than one perceptual system). The experiments reported here are preliminary empirical investigations into whether these two variables are involved in perceptually timing simple interceptive actions and whether perceptual systems other than vision are involved in time-to-contact perception.

To investigate these questions the type of experimental set up described in chapter 5 and which had been previously explored by Rosengren, Pick and von Hofsten (1988) and by Whiting and colleagues (for review see Savelsbergh, Whiting & Bootsma, 1989) was used. This set up involved the interception of self luminous targets in darkness. Rosengren et al. projected luminous tennis balls across a dark room at standing subjects who were required to catch these balls with one or with two hands. They found that subjects were able to catch luminous balls moving at speeds of roughly 8.6ms^{-1} or 6.7ms^{-1} in completely dark surroundings with one hand (though not nearly as well as they could under fully illuminated conditions). The results indicated that vision of the ball alone is sufficient for one-handed catching though not at a high level of skill. Vision of the environment and of the catcher's body provided by full illumination improves performance markedly. The question thus arises as to what additional information is being used by the subject when the environment and the body are visible.

The results of experiments by Smyth and Marriott (1982), Fischman and Schn-

ieder (1985) and Diggles, Grabiner and Garhammer (1988) established that one-handed catching is worse if the catching hand cannot be seen by the catcher. This decrease in the number of catches made when vision of the catching hand was occluded was not as great as the decrease found by Rosengren et al. This suggests that both vision of the hand and vision of the surroundings supply information used in one handed catching. Rosengren et al. attempted to assess the relative importance of vision of the hand and vision of the surroundings by selectively providing subjects with vision of the hand and of parts of the surroundings. To provide vision of the hand, subjects wore a glove with spots of luminous paint at the wrist and finger joints. To provide some information about the surroundings, six strips of luminous tape 2.54cm long were stuck to the wall either side of the ball's point of projection (at about 4.5 metres from the subject). It was found that while information about the surroundings in the form of the strips improved performance, vision of the hand did not. This is in apparent contradiction to the results mentioned above indicating that vision of the hand is useful in catching — catching performance in these studies was found to be worse when vision of the catching limb was prevented in a fully illuminated environment. Rosengren et al. found no difference between catching performance when the hand could be seen (luminous spots on the subject's glove) and when it could not.

The experimental situation of Rosengren et al. was not equivalent to that of Smyth and Marriott (1982). In the Smyth and Marriott paradigm, performance of subjects catching balls under conditions where all the information normally available to a subject is present is compared with their performance when some of this information is removed (vision of the catching limb, vision of the last 80-90 milliseconds of the ball's flight and vision of some of the environment behind the catching limb). Under the conditions of unrestricted vision, performance of subjects in all the published studies using the Smyth and Marriott paradigm is at or close to 100% and inter-subject variability is small. When vision of the catching limb is prevented there is a small decrement in performance and inter-subject variability remains low. Such a difference proves to be statistically detectable with relatively few subjects doing relatively few trials (fewer than 20 subjects performing ten trials or thereabouts in each condition in the studies mentioned). In the Rosengren et al. study, however, subject performance was poor in both conditions (hand visible, hand not visible) and inter-subject variability was relatively large. A small differ-

ence in performance is unlikely to be detectable without a large number of subjects and/or a large number of trials. Since Rosengren et al. used twelve subjects who performed only twelve catches in each condition their results are compatible with a small effect of seeing the catching limb. It should also be noted that the difference between the two conditions is "greater" in the Smyth and Marriott paradigm than in the Rosengren et al. paradigm: in the latter it is only vision of the hand that differentiates the two conditions, in the former part of the trajectory of the ball and part of the background are also involved. A greater difference in performance in the Smyth and Marriott paradigm might be expected for this reason. It cannot, therefore, be concluded from Rosengren et al.'s results that vision of the hand does not improve one-handed catching performance; it is clearly not vitally important, however, since any effects are rather small.

Rosengren et al. asked why the minimal visual information about the surroundings provided by the luminous strips should improve catching performance. They investigated the hypothesis that postural sway might interfere with catching and that providing information about the surroundings is able to reduce sway and so improve performance in the catching task. (It should be noted that this was an *post hoc* hypothesis since there was no *a priori* reason for supposing that postural sway would interfere with catching). The results showed that although sway increased in the dark (confirming previous work, e.g., Dichgans, Mauritz Allum & Brandt, 1976) the luminous strips did not reduce sway and no evidence was found to support the hypothesis that the luminous strips improved performance because they reduced sway. The position of the luminous strips in the frontal visual field means that they were unlikely to have provided much visual information about the small forward and backward motions of the head that are produced by postural sway (cf Anderson, 1986; Anderson & Braunstein, 1985). The fact that postural sway was not affected by the presence of the strips is consistent with this idea.

An alternative reason for the effect of the luminous strips is that their retinal motion is providing information about how the eyes are rotating relative to the environment, not about how they were translating. The time-to-contact information described in equations 4.21 and 4.23 in chapter 4 involves detecting the target object's rate of change of direction relative to a fixed frame of reference. As discussed in chapter 5, if the eye is tracking a target object this rate of change of direction is given by the rate of rotation of the eyes relative to the environment. The presence

of visible features fixed to the environment provides visual information about the target's rate of change of direction (see chapters 4 and 5 for detailed discussion). It was argued in chapter 5 that when only the target is visible, as in Rosengren et al.'s experiment, an observer could compute the target's rate of change of direction by tracking the target with pursuit eye-movements and measuring the rate of rotation of the eyes in the head and of the head relative to the environment using articular and vestibular proprioception to arrive at the rotation of the eyes relative to the environment. Such a route, though possible, is likely to be far less accurate than more direct visual information (it involves measuring the rate of change of target direction in a multimodal fashion and integrating proprioception from a variety of different sources). Visually one need only measure the rotational velocities of the retinal images of visual features which are fixed to the environment (see chapter 4 and Longuet-Higgins & Prazdny, 1980). Of course, when the environment is not visible because it is not illuminated, the rate of change of target direction cannot be measured visually and must be obtained from proprioceptive sources. It is to be expected, therefore, that under such conditions detection of a target's rate of change of direction will be degraded relative to its detection when visual information from the environment is available. Thus, the minimal environmental information provided by the luminous strips could be providing information about the target's rate of change of direction, which is important in the extraction of the kind of time-to-contact information that is likely to be used in the one-handed catching task reported by Rosengren et al. (1988).

§7.2. INVESTIGATION OF THE EFFECT OF VISIBLE SURROUNDINGS ON INTERCEPTIVE TIMING

Research indicates that detection of motion of the eye relative to the environment is most sensitive in the peripheral retina (Anderson, 1986; Boulton, 1988; Howard, 1982; Johansson, 1977). Thus, supplying a subject with minimal visual information about the fixed environment in the peripheral visual field should improve catching performance relative to performance when only the ball is visible if the interpretation of the luminous strip effect of Rosengren et al. (1988) given in the last section is valid. Experiment 1 explores this expectation using a similar task and experimental set-up to that employed by Rosengren et al.

Experiment 1

Methods

Subjects Eight subjects (undergraduate and graduate students and research staff at the university of Edinburgh) participated voluntarily. Five were male and three female, the median age of the subjects was 22.5 years and the range was 21-31 years (to the nearest year). All had normal vision or corrected to normal vision. None were informed of the hypothesis being tested and all stated that their preferred hand for catching and hitting was the right.

Apparatus Subjects were seated on a chair in a blacked out room lit with fluorescent strip lighting on the ceiling. They sat between two screens hung with black cloth on top of which were mounted two cardboard panels behind which black cloth was also hung, see figure 7.1 (black cloth was also hung behind the subject). Each cardboard panel was 50 × 75cm and had fifteen light emitting diodes (LEDs) on it: one in each corner and the rest placed within the rectangle formed by the four corner LEDs in a more or less random fashion. The LEDs emitted light of a greenish hue.

Subjects had four LEDs attached to the backs of their catching (right) hands just proximal to the knuckles. Light, hollow rubber balls (diameter 6 centimetres) were painted with self-luminous paint and then covered with clear transparent film (the balls were heated in an oven and then covered in film which then melted slightly and bonded firmly to the ball). The balls were then bound in clear plastic adhesive tape. These measures were necessary to prevent the paint from flaking off. These balls were thrown using a catapult constructed as follows: a commercially available “powerband” catapult band used for sports catapults was mounted on a wooden platform the inclination of which could be varied. The path of the ball over the platform was guided by a piece of plastic guttering (see figure 7.2)

The catapult was tested by using it to project balls at a large piece of card held where the subjects were to be seated. The inclination of the catapult was adjusted such that the balls thrown tended to arrive at or slightly below shoulder height. When a satisfactory inclination had been found a number of balls were projected at the piece of card which had been marked with circles of different radii (25cm, 40cm, 55cm) to get an idea of the variability in the paths of the balls thrown by the catapult. Fifty balls hit the card out of 54 thrown. Of the four that missed two failed to reach the card and two went too high. In the experiment trials on which

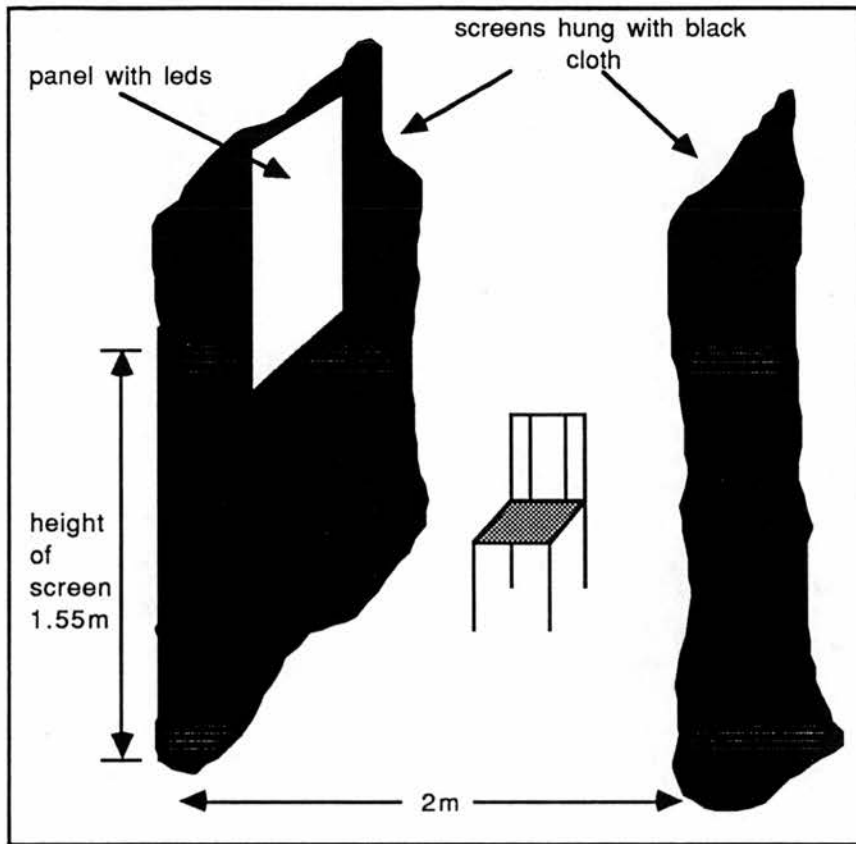


Figure 7.1. Subject was seated on a chair placed between two screens hung with black cloth. Black cloth was also hung over the door to the laboratory (not shown) which was about four feet behind the subject's chair. Panels were mounted on top of the screens on either side of the subject starting at about the subject's shoulder height.

balls failed to reach the subject were to be rejected as were those that went too high or too wide to be reached without the subject leaving the chair. 47 of the balls hitting the card landed within the circle of radius 55 cm; 35 landed within the 40 cm circle and 28 within the 25cm circle. Most of the variation in the position at which the balls struck the card was in the vertical dimension, however; 44 balls struck within the 25cm circle in the horizontal dimension. The catapult was tested once more for consistency after the experiment had been completed. Its performance was found to be more or less the same as it had been before the experiment was run. Adjustments to the inclination of the catapult were necessary from time to time and were made during practice sessions. The panel of LEDs on the subject's right was placed such that during a catching movement the hand and/or arm did not pass between the panel and the subject's eye(s).

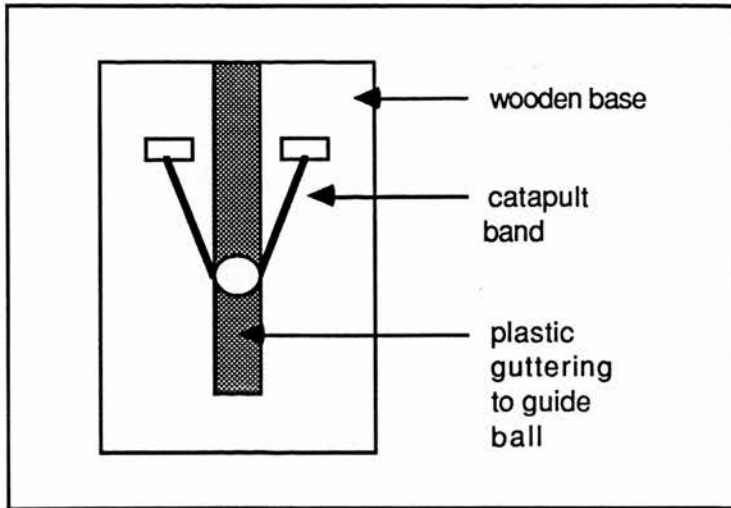


Figure 7.2. Diagram showing the top view of the catapult.

The horizontal speed of the balls was assessed roughly by measuring how long it took them to travel from the catapult to the card held at the point where subjects were to be seated, a distance of about 7 metres. This was done from a video with synchronized timing of ten projections which hit the card. The timing was accurate to within two video frame times (i.e., 80 milliseconds). Using this method the average horizontal ball speed was estimated to be 5.7ms^{-1} .

Procedure Subjects were seated and were required to catch balls with one hand projected at them (if the subject used two hands trials were re-run). Balls were projected to the subject's right and the subject seated such that the centre of the 50cm width in which most balls were found to fall was about 35-40cms to the right of the subject's shoulder. Subjects' performance was scored as follows: three points were awarded for a clean catch; two points for a fumbled catch, catch against the body or a catch subsequently dropped; one point for a touch of the hand and no points if the ball completely missed the hand.

It was found in pilot work that subjects tended to be very poor in the dark at first but were quick to improve. Pilot subjects reported that they began with considerable trepidation when catching in the dark and did not believe that they would be able to do the task at all but that they quickly gained confidence. To thoroughly familiarize experimental subjects with the task they were given two practice sessions, one on the day preceding and a second immediately before the experimental session. Practice sessions on the preceding day consisted of trials run

in full room lighting conditions with two eyes and with one eye (the left) occluded and in all the dark conditions: (1) ball only visible; (2) ball and hand visible (LEDs on the back of the catching hand were lit); (3) ball and peripheral lights visible; (4) ball, hand and peripheral lights visible. Ten practice trials in each condition were provided, five with two eyes open and five with the left eye occluded. Before the experimental trials were run subjects received a practice sessions consisting of six trials in full lighting conditions (three with binocular and three with monocular viewing) followed by six trials in each experimental condition (three binocular and three monocular in each) given in pseudo-random order. The intensity of the LEDs was adjusted using a variable power supply such that they were just visible to subjects in the periphery of their visual fields, if any LED was directly foveated it would become immediately invisible at this intensity. A final adjustment was made just before the experimental trials were run when the subjects could be considered to be more or less completely dark adapted.

Twenty four experimental trials were run in each of the following five conditions: (i) full room lighting; (ii) ball only visible (b); (iii) ball and hand visible (bh); (iv) ball and peripheral lights visible (bp); (v) ball, hand and peripheral lights visible (bph). Twelve of each set of twenty four trials were performed with binocular vision and six with monocular vision. The full room lighting conditions (excluded from the statistical analysis) were run before the conditions run in the dark. Subjects performed twelve trials binocularly followed by twelve monocularly (or *vice-versa*).

Trials in the other conditions were run in a counterbalanced order to minimize order effects. The following order permutations of four conditions ensure that each condition appears at each position in the order:

| | | | | |
|---------|-------|--------|--------|---------|
| orderA | 1 (b) | 2 (bh) | 4 (bp) | 3 (bph) |
| order B | 2 | 3 | 1 | 4 |
| order C | 3 | 4 | 2 | 1 |
| order D | 4 | 1 | 3 | 2 |

Each subject received one of these orders of conditions (each order was received by two subjects). Experimental sessions were divided into two groups of 48 trials separated by a short break of three or four minutes. If the conditions were ordered according to, for example, order A during the first group of trials they would be presented in reverse order in the second group. Twelve trials were run

consecutively in each condition, six binocularly followed by six monocularly or vice versa (this order was determined pseudorandomly). The following table illustrates a distribution of trials that might occur using this design (order A is assumed for the sake of illustration):

| | | | | | | | | |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| trial group | | | 1 | | 2 | | | |
| condition | b | bh | bp | bph | bph | bp | bh | b |
| number of eye | 1 2 | 2 1 | 2 1 | 1 2 | 2 1 | 2 1 | 1 2 | 2 1 |
| number of trials | 6 6 | 6 6 | 6 6 | 6 6 | 6 6 | 6 6 | 6 6 | 6 6 |
| total trials | | | 48 | | 48 | | | |

Results

The total scores for the subjects in the various conditions are presented in table 7.1. The mean scores are plotted in figure 7.3. Four points may be noted from the graph: (1) Performance in full lighting conditions is much better than in any of the dark conditions with six out of the eight subjects catching every ball cleanly in the binocular full room lighting condition. (2) Binocular viewing results in considerably better performance than monocular viewing in all conditions. (3) Performance with the peripheral LEDs visible appears to be better than performance without them. (4) Vision of the LEDs on the hand appears to make no difference to performance. The same picture is repeated in figures 7.4(a) and (b) which show the mean numbers of clean catches and complete misses respectively. These indices of performance show the same pattern as figure 7.3: catches are more frequent and misses less frequent (1) in the light as compared with the dark conditions, (2) under binocular viewing conditions compared to monocular viewing conditions and in the light, and (3) when the peripheral lights were visible compared with when they were not.

A statistical analysis of the scores data was conducted on the conditions run in the dark. A 2×2 repeated measures analysis of variance (ANOVA) with multiple comparisons was used. Both main effects were highly statistically significant as might be expected simply from inspection of the data in table 7.1. Binocular viewing was much better than monocular viewing ($df= 1, 7; F = 295.3; p < 0.0001$) and the lighting conditions affected performance ($df= 3, 21; F = 33; p < 0.0001$). There was no significant interaction between number of eyes used and the lighting conditions ($df= 3, 21; F = 0.74; ns$).

| condition | Monocular | | | | | Binocular | | | | |
|-----------|-----------|------|------|------|-------|-----------|------|------|------|-------|
| | Dark | Hand | Peri | both | light | Dark | Hand | Peri | both | light |
| subject 1 | 9 | 12 | 10 | 13 | 26 | 16 | 15 | 22 | 17 | 36 |
| 2 | 10 | 12 | 11 | 12 | 23 | 17 | 11 | 20 | 19 | 34 |
| 3 | 12 | 13 | 17 | 17 | 34 | 18 | 19 | 23 | 27 | 36 |
| 4 | 20 | 19 | 22 | 23 | 31 | 28 | 27 | 33 | 33 | 36 |
| 5 | 16 | 15 | 19 | 24 | 33 | 26 | 22 | 30 | 30 | 36 |
| 6 | 12 | 8 | 19 | 16 | 23 | 23 | 24 | 27 | 26 | 36 |
| 7 | 15 | 17 | 19 | 19 | 30 | 20 | 20 | 29 | 27 | 36 |
| 8 | 19 | 18 | 20 | 21 | 27 | 26 | 23 | 26 | 32 | 34 |

Table 7.1. Total scores for the eight subjects in experiment 1 in the five different illumination conditions using either one or two eyes. Key: Dark = ball only visible; Hand = LEDs on hand visible; Peri = LEDs on the peripheral panels visible; both = both panel and hand LEDs visible; light = full room lighting.

In order to examine statistically (1) whether vision of LEDs on the hand made a difference to catching performance and (2) whether peripheral lights made a difference to catching performance paired comparisons between the means in various conditions are required. Out of a total of seven possible mutually orthogonal planned comparisons, three were of interest: (1) the two ball only visible conditions compared with the two hand and ball visible conditions; (2) the two monocular conditions in which the peripheral LEDs were not visible compared to the two monocular conditions in which they were visible: (3) same as (2) for the binocular conditions. The significance of these comparisons was assessed using Scheffé's test (Scheffé, 1953). The value of α is set to 0.1 (rather than 0.05) when this test is used for planned comparisons (Scheffé, 1953). The value of F computed by Scheffé's method is compared with $F' = (k-1)F$, where F' is the tabled value of F for $k-1$ and $k(n-1)$ degrees of freedom (k is the number of conditions and n the

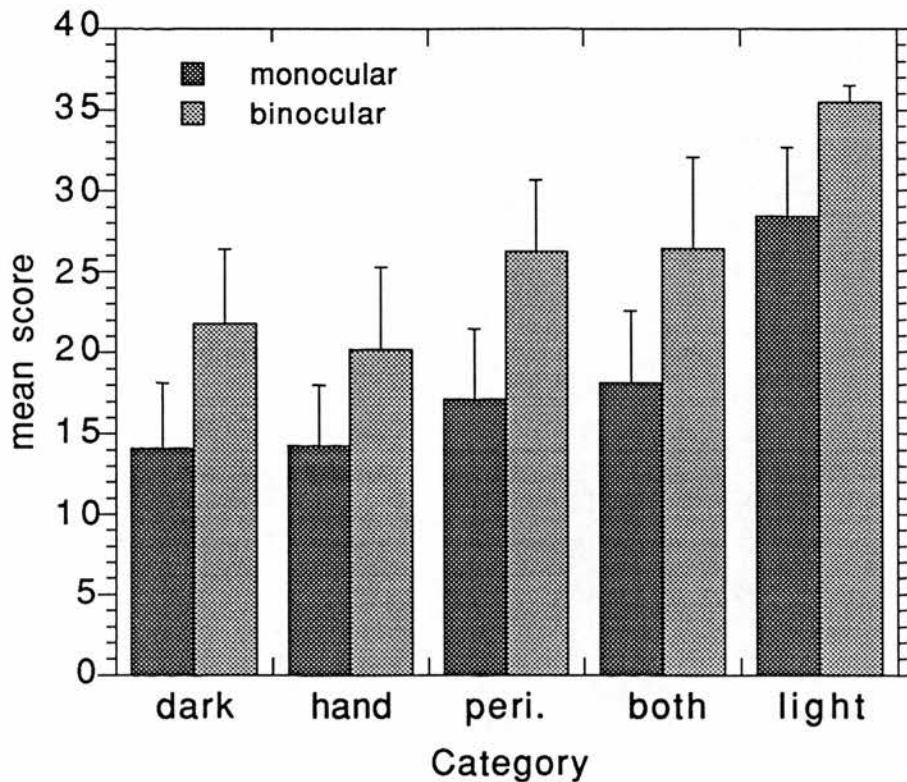


Figure 7.3. Mean scores across subjects in experiment 1. Error bars indicate the standard deviation of the scores. Key as for table 7.1.

number of observations in each condition). It was found that the LEDs on the hand had no effect on performance ($F_{7,56} = 1.03$, ns.). The peripheral LEDs, however, significantly improved performance in both monocular ($F_{7,56} = 21.7$, $p < 0.1$) and binocular conditions ($F_{7,56} = 53$, $p < 0.05$).

Discussion

The results were consistent with other reports (Savelsbergh et al., 1989; Rosengren et al., 1988) that minimal information about the fixed environment is sufficient to improve performance in a one-handed catching task relative to performance when only the ball to be caught is visible. The difference here is that vision of the environment was confined to the peripheral visual field and present at the edges of the field. In the other studies vision of the environment was confined to the central visual field. The results also confirm the finding of Rosengren et al. that vision of lights attached to the catching hand have no discernable effect on catching performance. Further, the results show that it is possible to catch a self-

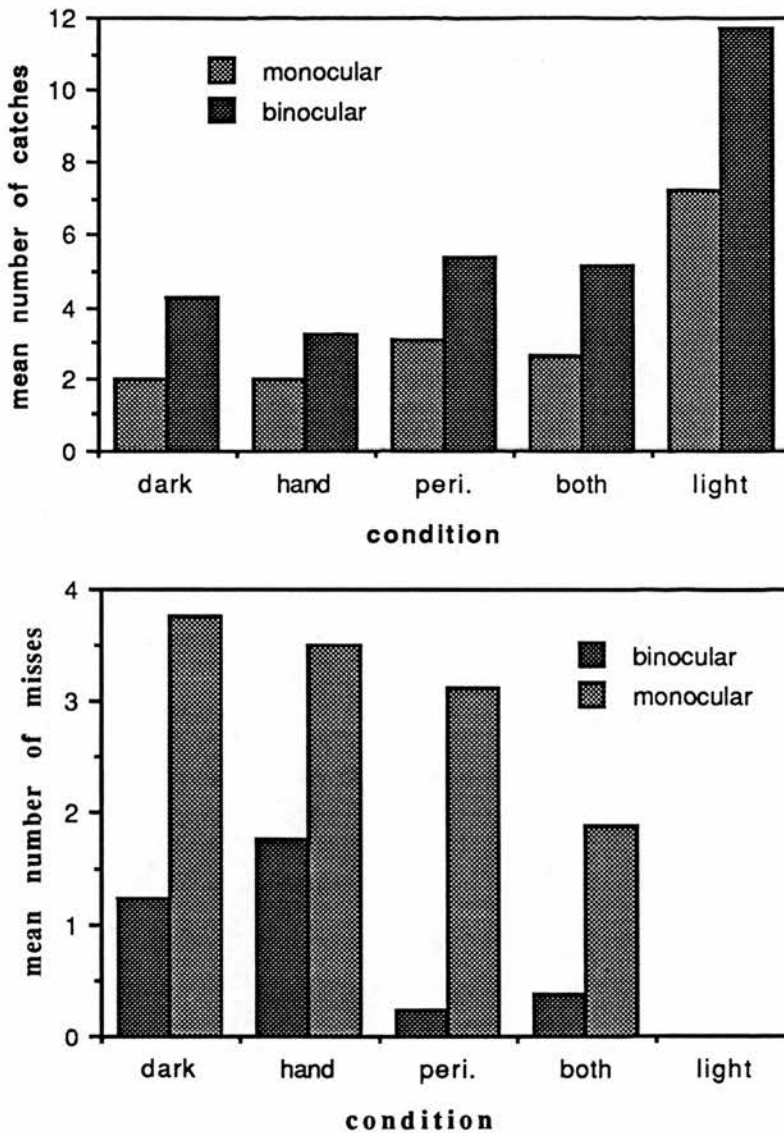


Figure 7.4 a) Mean number of catches (out of a possible maximum of 12) across subjects. **b)** Mean number of misses (out of a possible maximum of 12); no misses were recorded in the binocular full lighting condition and only one in the monocular full lighting condition.

luminous ball in the dark when only one eye is being used, though performance is considerably degraded compared to that with two eyes.

The insignificance of the effect of LEDs on the catching hand cannot be used to argue that visual information about the catching limb does not influence catching performance. What seems to be implied by the nonsignificant results reported here and by Rosengren et al. is that vision of the catching limb has at most a small effect on performance which, if present, cannot be detected by the relatively

crude scoring measures used. The results of Smyth and Marriott (1982) showing a significant detrimental effect of occlusion of the catching limb may have been due, in part, to (a) occlusion of the ball during the last part of its trajectory (about 90ms) and/or (b) occlusion of part of the fixed environment on the subject's catching side. In order to determine whether vision of the catching limb *per se* has an effect on catching performance one would employ the luminous target in the dark paradigm and use more precise measures of performance. Experiments 4 and 5 reported below examine one aspect of this question in the special case of a pure timing task.

The pertinent finding in this experiment is the effect of minimal peripheral vision of the surroundings on performance. As predicted, catching performance was improved by the presence of these lights and this effect is present in both the monocular and binocular conditions. This result is consistent with information about how the eyes are moving relative to the environment being used in catching and hence with the use of the target's rate of change of direction (which is equal to the rate of rotation of the eyes relative to the environment when they are pursuing the target). This result is thus consistent with use of the time-to-contact information described in chapter 4 in timing catches. However, it does not establish that this information is used: indeed, from the results of this experiment it is not possible to determine whether the peripheral lights improve the positioning component of the catch or the timing component. It is, in principle, impossible to determine whether errors in positioning or timing are being affected since a given observed error cannot be identified as due to faulty positioning, faulty timing or some combination of the two (cf Lee, 1980a; chapter 3), even if one had precise recordings of the movements made in attempting a catch. For example, if it were found that the fingers closed too much before the ball arrived such that it hit them, then although this might look like a timing error (the hand is on the path of the ball so positioning is in some sense correct) the timing of the finger closure would be correct for a slightly different position and hence one could treat the failure of the catch as due to a positioning error. There is no objective way to decide which interpretation of the failed catch is the correct one, even if one accepts that, in principle, such a decision is meaningful. It is worth noting that such a decision may not be meaningful in the context of questions concerning timing information. Suppose that time-to-contact information is used not only to control the timing of the catch, but also in computing the position where the catch should take place

If time-to-contact information is used in computing position, this would mean that an error in estimating time-to-contact would lead to both timing and positioning errors. For these reasons it was decided to investigate the role of peripheral visual information in a pure timing task with no positioning component.

Experiment 2a

As discussed earlier, if the time-to-contact information described in chapter 4 is used to time the interception of moving targets which bypass the observer then the rate of change of target direction relative to the point of observation is required. This implies that timing an interception of a self-luminous target should be better in the presence of peripheral visual information about the fixed environment than in darkness. If vision is the only perceptual system that provides the information for interceptive timing, then when only the moving target is visible timing should be based upon visual information provided by the target alone, i.e. local tau., and errors in timing should reflect this (see below). If information about rate of change of target direction can be extracted by articular proprioceptive systems then timing errors should reflect this. As argued earlier (see above and chapter 5) visual information is expected to be more accurate and less prone to noise than alternative articular proprioceptive sources; this means that timing is expected to be both more accurate and more consistent or precise (less variable) when visual information about the surroundings is present than when it is not. It is also expected that the less the moving target changes its direction relative to the observation point (and consequently, the less the eyes need to move to keep it foveated) the better timing should be when there is no visual information about the surroundings.

To test these predictions a ball-trapping task (described in detail below) involving no positioning movements or positional uncertainty was employed. To test the prediction that the more the eyes move relative to the environment during pursuit of the moving target, the worse timing will be when there is no vision of the surroundings, subjects performed the task in two positions, one requiring more movement of the eyes than the other.

Methods

Subjects Sixteen subjects took part voluntarily in this experiment. Four of them had taken part in experiment 1 but had remained ignorant of the experimental hypothesis. All were students and staff at the University of Edinburgh. 9 were male and 7 female. The median age was 23 years and the range 21-45 years. All had normal or corrected to normal vision and stated that their preferred hand for catching or hitting was the right.

Apparatus A ball trapping apparatus was constructed according to the design illustrated in figure 7.5 which indicates the dimensions. The last 90cms of the trackway (the trapping zone) lay under a sheet of rigid perspex (9mm thick) in the centre of which was mounted a handle which could be grasped by subjects. The perspex sheet was hinged along one edge and hung from a bar by springs on the opposite edge. When the sheet was hung by the springs a ball rolling along the trackway was able to pass under the sheet without touching it. The sheet could be pressed down and balls trapped beneath it. The springs were not very stiff and offered little resistance to movement of the sheet over the small distance necessary to trap a ball.

Self-luminous balls were rolled down the trackway. Two balls were used and were made as follows: two solid rubber balls of diameter 6cms were painted with self-luminous paint of the same kind as that used in experiment 1. The balls were then coated with a layer of commercially available silicon-rubber waterproof sealant to prevent the paint flaking off. This sealant did not smoothly cover the balls which had a slightly irregular surface after application of the sealant. The irregular surface meant that the ball's speed at any point on its roll down the trackway would not be the same from trial to trial even if rolled from exactly the same starting position. The balls' surfaces were smooth enough to prevent them bouncing about as they rolled down the trackway.

Balls were started rolling from one or other of two positions (20cms and 30cms) behind a black curtain which obscured from the subject's view the first few centimetres of their path. The distance from the point where the balls came into view to the trapping point under the handle on the perspex sheet was 1.95 metres.

The ball's speed before entry to the trapping zone could be estimated using the time taken for the ball to pass through a pair of infra red "light-gates" placed as illustrated in figure 7.5. Each light-gate consisted of an infra red emitting LED mounted just above the edge of the trackway opposite an infra red sensor. These two light gates were 20cm (± 1 mm) apart and connected to a Bassin millisecond electronic timer. The time taken for the ball to pass between the two light gates could thus be recorded; the speed of the ball at entry to the trapping zone may then be estimated to be equal to (distance between light gates) \div (time for ball to pass between light gates).

The experiment was conducted in a small blacked-out room (lighting was pro-

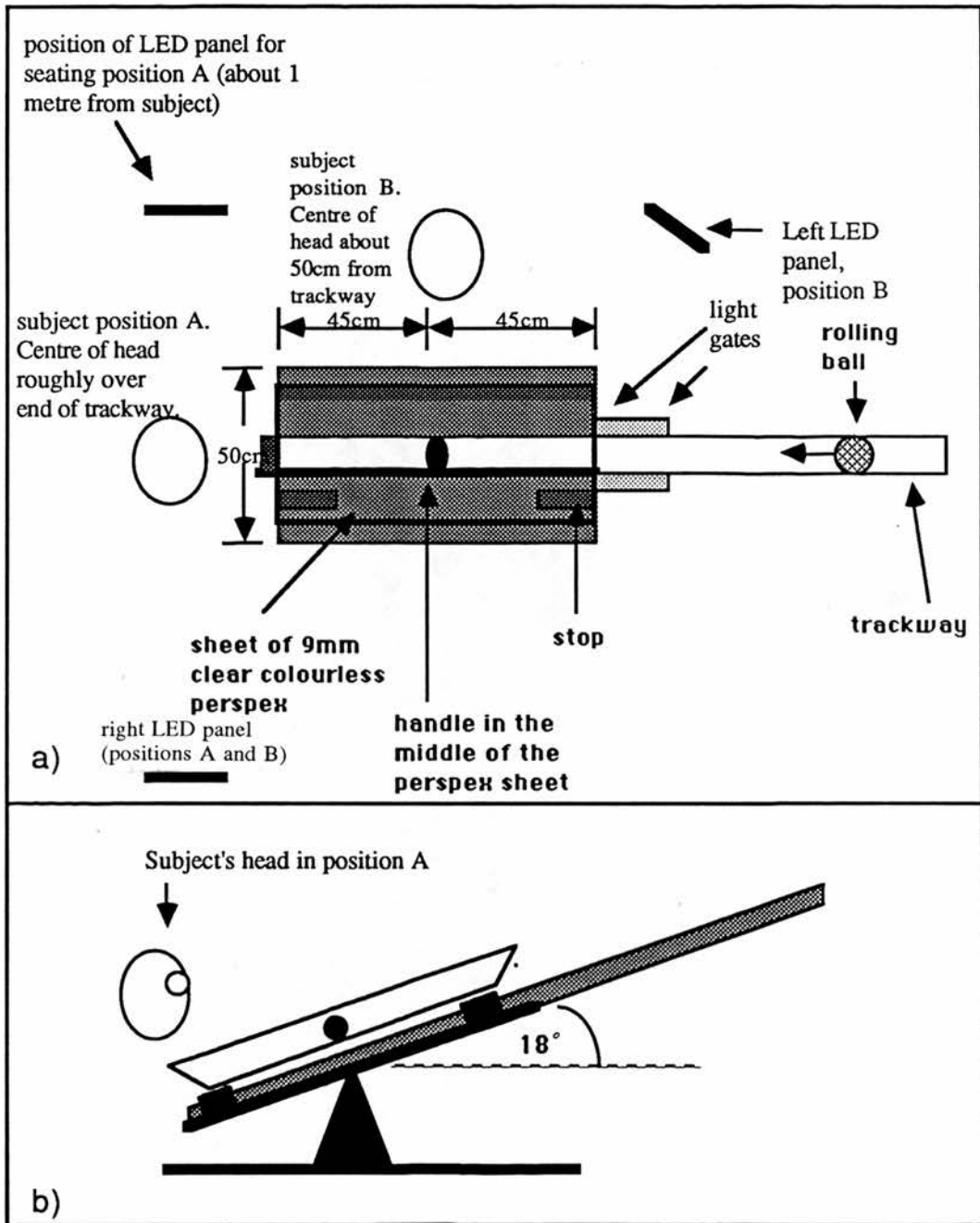


Figure 7.5 Schematic diagram of the ball-trapping apparatus. (a) Plan view, (b) side view.

vided by two strip lights on the ceiling), the walls of which were hung with black cloth.

Procedure. The subjects' task was to trap balls rolling along the trackway directly

under the hand gripping the handle. If this was not achieved the ball would be trapped under the perspex sheet some distance from the handle. The handle had an opaque circular base with a diameter of 6cms. There were two independent variables (subject position and lighting condition) each with two levels. There were thus four different conditions: (1) subject seated in position A and only the ball visible (DA); (2) subject seated in position A with ball and peripheral lights visible (PA); (3) subject seated in position B with only the ball visible (DB); (4) subject seated in position B with ball and peripheral lights visible (PB). Fourteen experimental trials were run under each condition. The position of the subject in each positions A and B are indicated in figure 7.5, the positions of the LED panels were adjusted when the subject changed seating position (see figure 7.5).

Before the experimental trials were run subjects received two practice sessions. Pilot work indicated that subjects required some experience with the task before they began to perform successfully. On the day before the experimental trials were run subjects were thoroughly familiarized with the task. They first performed it under full room lighting conditions and received as many trials as it took for them to trap five balls in succession directly under the base of the handle in both seating positions. When this had been achieved¹ subjects were given ten practice trials in each of the experimental conditions (five in each seating position). Subjects received the second practice session immediately before the experimental trials were run. First, twenty trials in full room lighting were run (ten in each position). Next, six trials (three in each position) in each of the four experimental conditions were run in a pseudorandom order. After a short break experimental trials were run, by which time subjects had been in the dark for about twenty minutes. The intensity of the LEDs was adjusted using a variable power supply such that they were just visible to subjects in the periphery of their visual fields, if any LED was directly foveated it would become immediately invisible at this intensity.

A similar method of counterbalancing to that used in experiment 1 was adopted. Four orderings of conditions were chosen in the same way as in experiment 1. Subjects received seven trials in each condition ordered according to one of the four orderings. There was then a short break of two to three minutes followed by seven trials in each condition this time given in the reverse order.

¹ Subjects required on average 33 trials to make five traps in a row. One person did not achieve this in a reasonable length of time and did not look as if she would achieve it: she was not used as an experimental subject.

Errors were measured as follows. Figure 7.6 shows the top of the perspex sheet defining the trapping zone. The distance of the ball's centre from the edge of the handle base was measured by placing a clear rigid plastic strip with a line drawn down its centre over the ball and reading off the distance in centimetres where this line crossed the tape measure (the strip had the same width as the ball). Using the speed at entry to the trapping zone the error in timing a ball trap was estimated by dividing the distance measured as above by the speed of the ball. Subjects' errors were summarised by the mean of the signed errors (called the constant error) and their standard deviation (called the variable error).

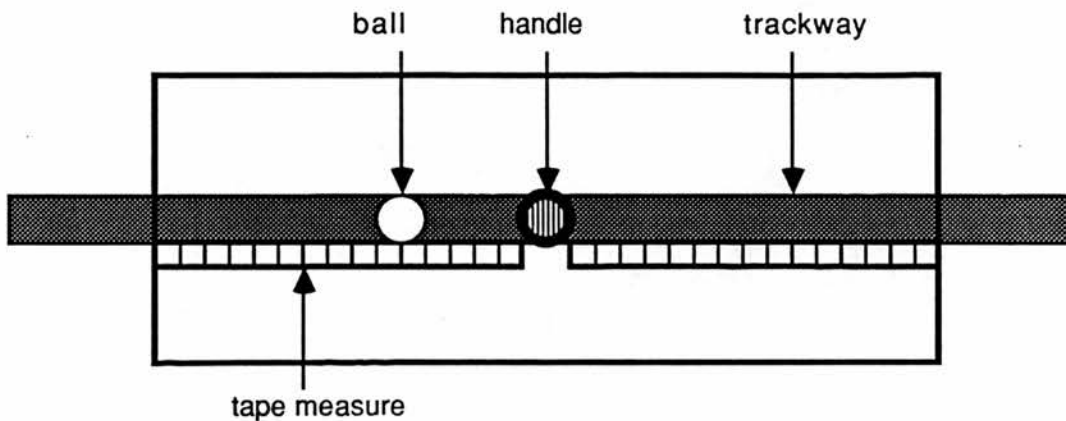


Figure 7.6 Measurement of ball position error. The distance of the centre of the trapped ball from the edge of the handle base could be read off from the tape measure attached to the perspex sheet.

Results

Mean variable errors (VE) and constant errors (CE) over subjects are shown in figures 7.7 and 7.8 and the descriptive statistics are tabulated in table 7.2a and b. Both sets of data show a similar pattern: errors are larger when the peripheral lights are not present and the effect of the peripheral lights is greater in position B than in position A as predicted. This latter interaction effect is shown more clearly in figures 7.9 and 7.10. Analyses were conducted to assess the statistical significance of these effects.

The CE data departs strongly from the normal distribution (see table 7.2a), the values for kurtosis and skewness are too great for a parametric statistical analysis. No simple transformation of the data was found that could significantly improve

a)

| | DA | PA | DB | PB |
|-----------------|-----|-----|------|-----|
| mean | 9.2 | 5.9 | 10.7 | 5.4 |
| sd | 6.3 | 5.3 | 10 | 6.7 |
| skewness | 1.3 | 1.1 | 1.9 | 2.5 |
| kurtosis | 1.6 | 0.2 | 3.6 | 5.8 |

b)

| | PA | DA | DB | PB |
|-----------------|------|------|------|------|
| mean | 23.8 | 19.2 | 29.7 | 19.6 |
| sd | 8.1 | 4.3 | 8.0 | 5.3 |
| skewness | 0.7 | 0.5 | 1.0 | 0.2 |
| kurtosis | -0.3 | -0.9 | 0.5 | -0.8 |

Table 7.2. Descriptive statistics for a) the unsigned constant errors and b) the variable errors (in milliseconds)

this state of affairs. For these reasons Wilcoxon signed ranks tests were used to analyze the CE data and thus the statistical significance of the interaction could not be assessed. The sign of the CE was not important here and the analysis was conducted on the magnitude of the CEs. It was predicted that if there were an effect of peripheral lights on the CE (accuracy) of interceptive timing then it would tend to reduce the error (improve accuracy), thus a one-tailed test of significance was used. It was found that errors in the presence of peripheral lights were significantly smaller than when only the ball was visible in both position A ($T = 32, p < 0.05$) and position B ($T = 21, p < 0.01$). The results strongly suggest that the effect is more powerful in position B than in position A (see figure 7.10).

The VE data is closer to normal than the CE data and the values for kurtosis and skewness are within the acceptable range for parametric statistical analysis. There is, however, a tendency for the slightly larger standard deviations to be associated with the larger means (see table 7.2b) which violates the conditions for parametric statistics. This effect was found to be reduced by taking the natural logarithms of the VEs and this transformed data was analyzed using a two-way repeated measures ANOVA. It was found that there was a significant effect of illumination condition ($df = 1, 15; F = 16.2; p < 0.01$) and a significant interaction

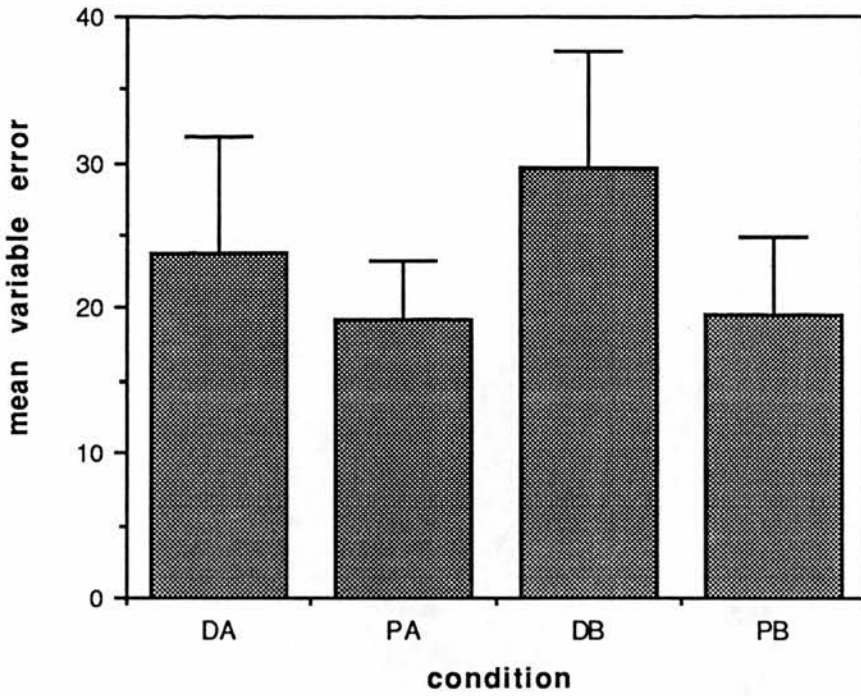


Figure 7.7 Bar chart of the mean variable errors in milliseconds across subjects in the four experimental conditions. Key: DA = ball only visible, subject in position A; PA = peripheral lights visible, subject in position A; DB = ball only visible, subject in position B; PB = peripheral lights visible, subject in position B.

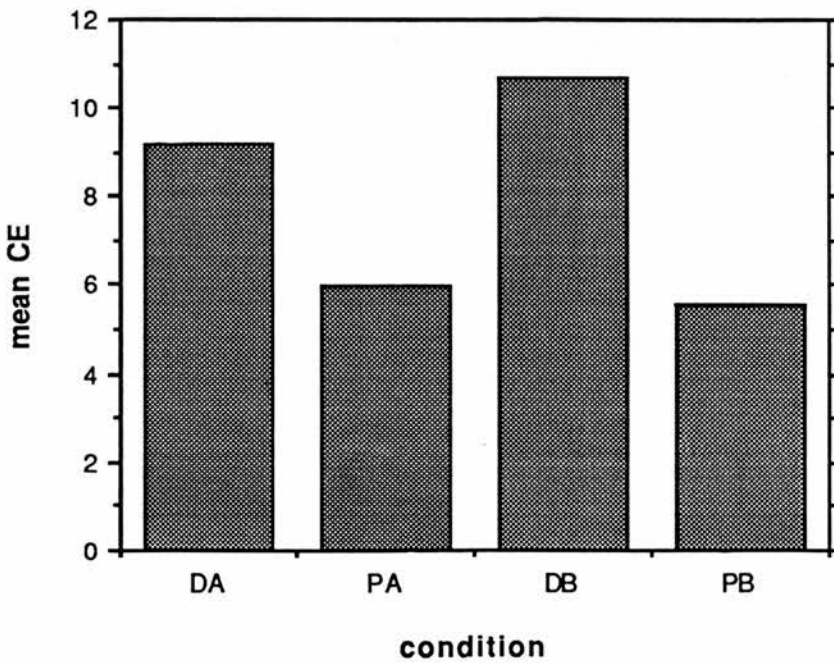


Figure 7.8. Bar chart of the mean unsigned constant errors (CE) in milliseconds across subjects. Key as for figure 7.7

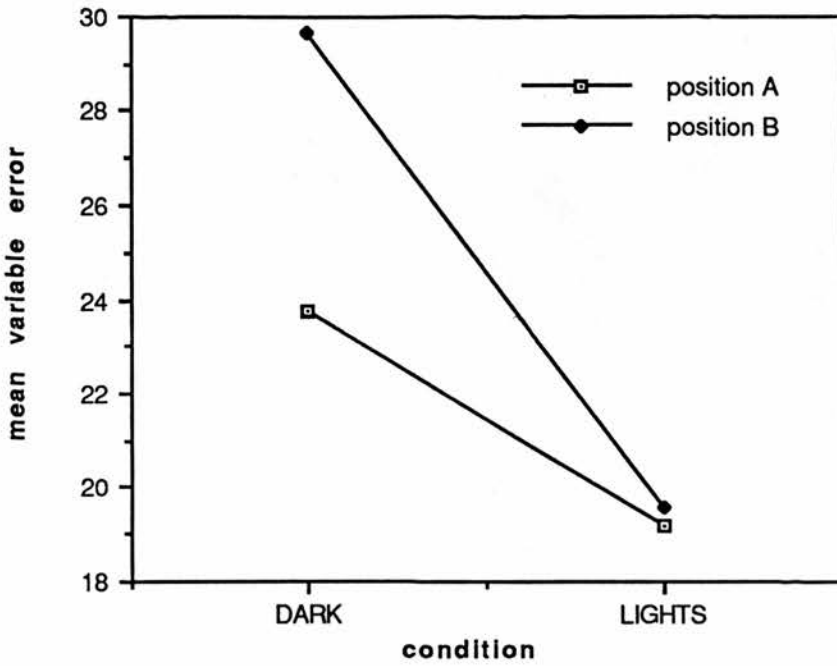


Figure 7.9. Graph showing the interaction between subject position (A or B) and lighting condition (DARK = ball only visible; LIGHTS = peripheral lights visible) for the mean variable error in milliseconds.

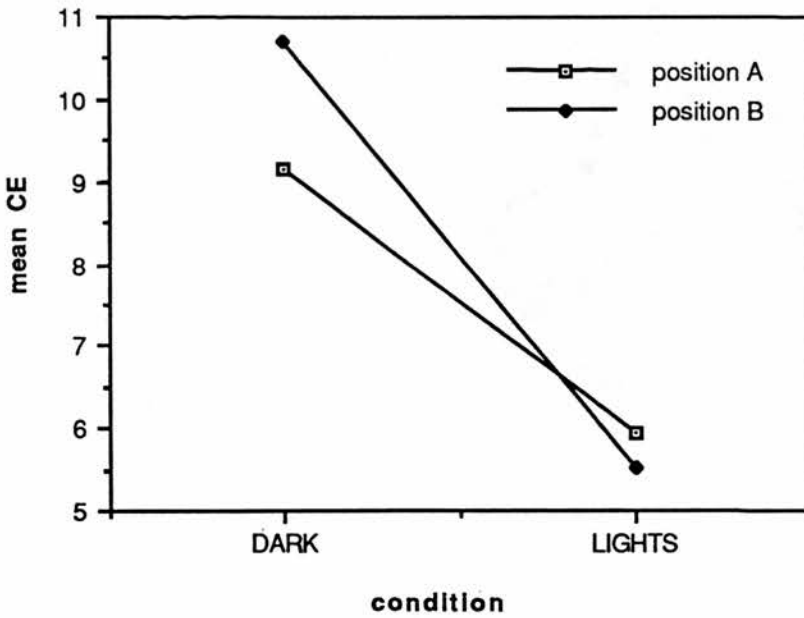


Figure 7.10. Graph showing the interaction between position and illumination condition (both as for graph figure 7.9) for the mean constant error in milliseconds.

between subject position and illumination condition ($df= 1, 15; F = 11; p < 0.01$) as predicted. There was no significant main effect of seating position ($df= 1, 15;$

$F = 1.65$; $p \gg 0.1$). Planned comparisons were made to analyze the interaction. Peripheral lights improved performance both in seating position A ($df = 1, 15$; $F = 7.79$; $p < 0.05$) and seating position B ($df = 1, 15$; $F = 54.5$; $p < 0.0001$) though the effect is considerably greater in position B as predicted. Further, VEs in seating position B were significantly greater than in position A when only the ball was visible ($df = 1, 15$; $F = 17.5$; $p < 0.01$) but there was no significant difference between the two seating positions when the peripheral lights were present ($df = 1, 15$; $F = 0.001$; ns). These effects are shown clearly in figure 7.9.

Timing error was also correlated with ball speed to investigate whether there was any relationship between the two (for reasons discussed below). No relationship was found: the correlations and r^2 values for the individual subjects are listed in table 7.3.

| subject no. | correlation coefficient | r^2 |
|-------------|-------------------------|-------|
| 1 | -0.296 | 0.088 |
| 2 | -0.159 | 0.025 |
| 3 | 0.053 | 0.003 |
| 4 | -0.082 | 0.007 |
| 5 | 0.106 | 0.011 |
| 6 | 0.222 | 0.049 |
| 7 | -0.076 | 0.006 |
| 8 | 0.147 | 0.021 |
| 9 | 0.199 | 0.04 |
| 10 | 0.010 | 0.000 |
| 11 | -0.163 | 0.010 |
| 12 | 0.149 | 0.022 |
| 13 | 0.032 | 0.001 |
| 14 | -0.043 | 0.002 |
| 15 | -0.131 | 0.017 |
| 16 | -0.219 | 0.048 |

Table 7.3 Pearson coefficients of correlation between ball speed and timing error and associated r^2 values for all subjects in experiment 2a.

Discussion

The results support the predictions that both the precision (VE) and accuracy (CE) of interceptive timing would improve in the presence of peripheral visual in-

formation about the fixed environment and that this improvement would be more marked when more movement of the eyes was required to pursue the target. The presence of the peripheral lights improved performance in both seating positions, more so in position B than position A. With only the ball visible subject's performance in position A was significantly better than in position B, but when the peripheral lights were visible no difference in performance was found. If the peripheral lights are providing information about how the eyes are moving relative to the environment as the moving target is being pursued, exactly the pattern of results reported here is predicted as explained above. The results are therefore consistent with use of the target's rate of change of direction in interceptive timing and thus with the use of the time-to-contact information involving this variable (Lee & Young, 1985; chapter 4).

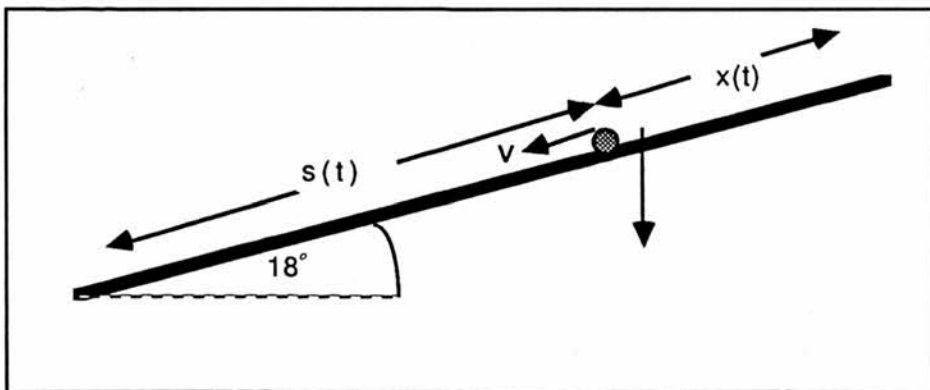


Figure 7.11. Model of experimental set-up. The trackway is assumed to exert negligible frictional force on the rolling ball which may be treated as a particle which has a velocity at time t of v m/s and an acceleration down the slope of $g\sin 18^\circ = 3.0$ m/s² (neglecting angular momentum).

The results are not what would be expected were local tau alone being used as the source of timing information in the experimental task and not only because such a strategy does not predict the effect of peripheral visual information but also because subjects' accuracy is simply far greater than one would expect from use of local tau alone. To see this consider the simplified model of the experimental set-up illustrated in figure 7.11. The actual time-to-contact, T_c , as a function of distance from the interception point, $s(t)$, may be derived as follows. From Newton's standard kinematic equations for the motion of a particle we have

$$s(t) = -v(t)T_c(t) - \frac{3.0T_c^2}{2}. \quad (7.1)$$

By using the standard formula for the roots of a quadratic, 7.1 can be solved for $T_c(t)$. We are interested in the positive root which is:

$$T_c(t) = \frac{-v(t) + \sqrt{v^2(t) + as(t)}}{a} \quad (7.2)$$

This can be written as a function of $s(t)$ alone by noting that $v^2(t) = 2ax(t)$ where $x(t)$ is the distance L from the starting point to the interception point less $s(t)$. Thus, $v(t) = \sqrt{2a(L - s(t))}$.

The time-to-contact assuming constant velocity, T_c^* is given by

$$T_c^*(t) = \frac{-s(t)}{v(t)} = \frac{s(t)}{\sqrt{2a(L - s(t))}}. \quad (7.3)$$

Finally local tau is given by equation 6.9 which may be written,

$$\tau_L(t) = T_c^*(t) + \frac{D^2}{[s(t)\sqrt{2a(L - s(t))}]}, \quad (7.4)$$

where D is the distance of the eye from the interception point (which is at the point of nearest approach to the eye). For illustrative purposes the following values for the parameters will be assumed: from the dimensions of the experimental apparatus as illustrated in figure 7.11, $a = 3.0\text{ms}^{-2}$, $D = 50\text{cm}$, and $L = 2.25\text{m}$ (this value for L means that balls enter the trapping zone at about 2.3ms^{-1} , roughly the average speed of entry in the experiment).

Figure 7.12a shows time-to-contact computed according to equations 7.2, 7.3 and 7.4 plotted as a function of distance from the interception point. It is clear that although the constant velocity strategy provides a very good estimate of the actual time-to-contact over the last 75cms or so, the utility of local tau begins to break down seriously over the final part of the trajectory. A constant velocity strategy would thus be effective in the experimental task. Figure 7.12b compares tau with time-to-contact when the target moves with a constant velocity. In figure 7.13a local tau is plotted as a function of time-to-contact (assuming constant target velocity) and distance (D) from the observation point of the point of nearest approach for a fixed target speed. The left hand visible face of the figure is the graph of local tau against time-to-contact for a direct collision approach ($D = 0$) and in this case local tau is equal to the time-to-contact. As D gets larger the error in taking local tau as an estimate of time-to-contact gets larger. In figure 7.13b local tau

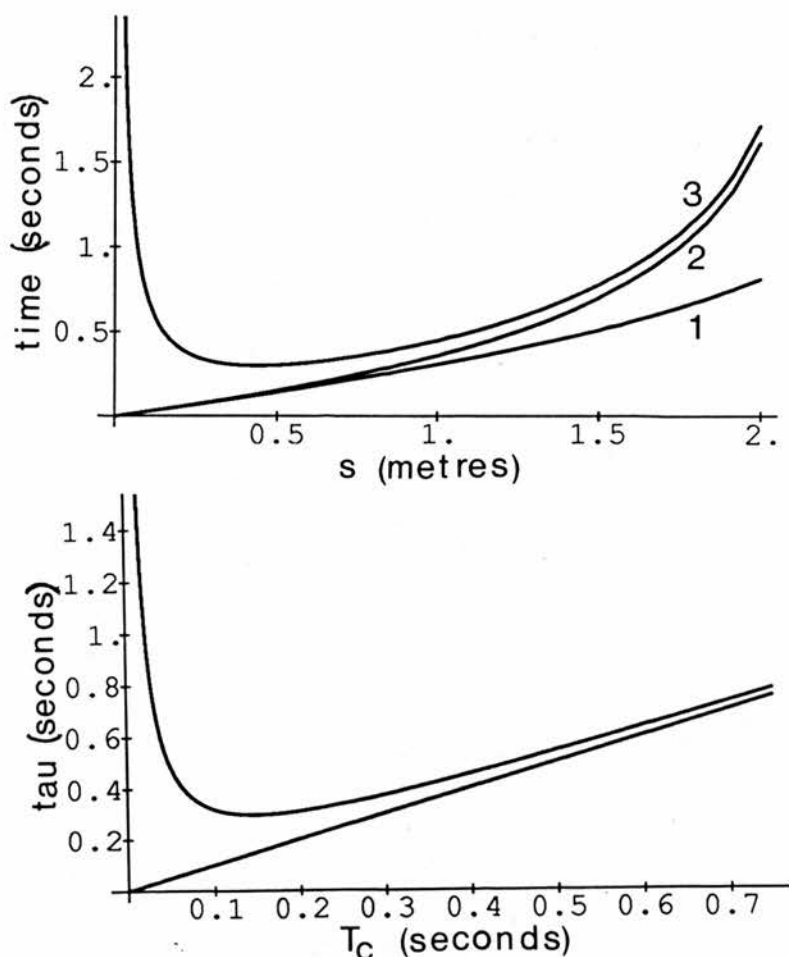


Figure 7.12 a) Time-to-contact computed according to three different methods plotted as a function of distance from the interception point. Key: 1) actual time-to-contact; 2) time-to-contact assuming the constant velocity strategy; 3) time-to-contact as estimated by local tau. b) Plot showing how local tau varies with time-to-contact under the assumption of constant velocity. The straight line is the plot that would result were tau equal to the time-to-contact.

is plotted as a function of time-to-contact (assuming constant target velocity) and target speed for a fixed value of D . The flat region of the graph is where tau provides a reasonable estimate of the time-to-contact. When the target speed is low and/or the time-to-contact is small the estimate provided by local tau diverges wildly from the actual time-to-contact.

As described in chapter 6, large timing errors result from using local tau as the source of timing information in tasks such as that used in this experiment where the target is moving relatively slowly. As can be seen clearly from the graphs (figures 7.12 and 7.13), when the constant velocity strategy begins to become accurate, use of local tau becomes increasingly inaccurate. Even where tau provides the best estimate of the actual time-to-contact (between about 0.7 and 1 metre before contact) the error is about 100ms (at 0.7m before contact $\tau_L \approx 330$ ms and

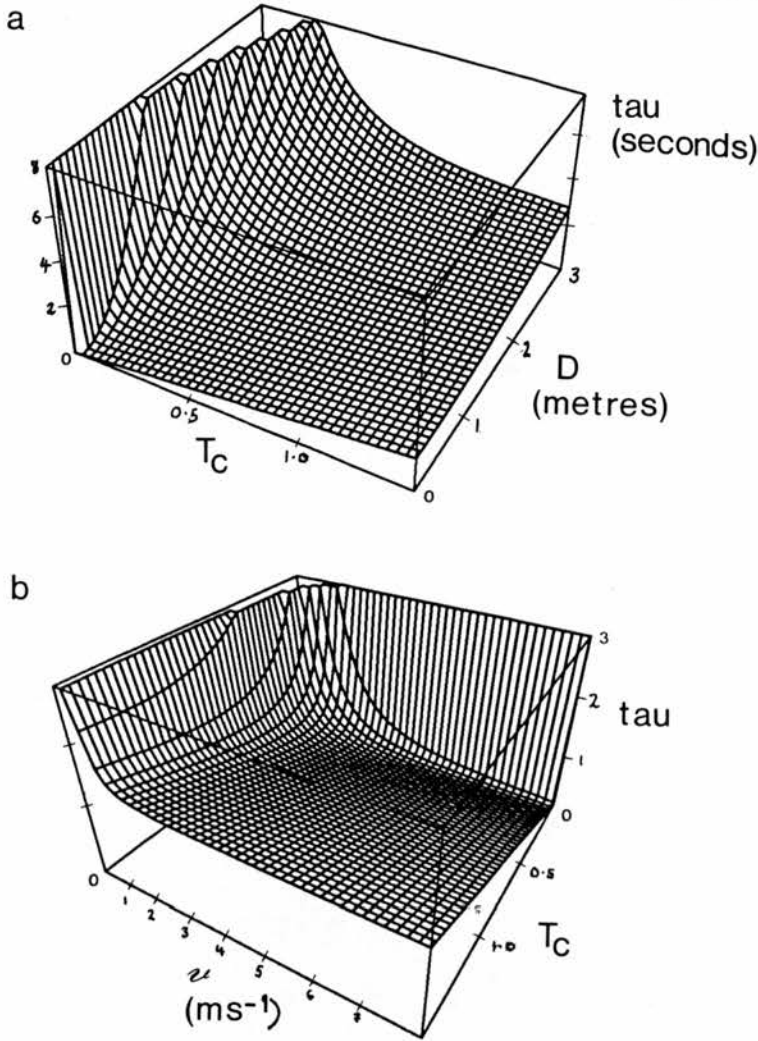


Figure 7.13 a) Tau plotted as a function of time-to-contact (T_c) and distance (D) of the point of nearest approach from the observation point for a particular target velocity (2 m/s). b) Tau plotted as a function of time-to-contact and target speed (v) for a particular distance, D (0.5 metres). Target velocity is considered constant in both cases.

$T_c \approx 200\text{ms}$). The mean timing errors in position B found in this experiment were quite small. The mean constant timing error magnitude with only the ball visible was 10.7ms and in other conditions the mean CE was smaller than this. These errors are thus much smaller than would be expected were subjects using local tau as information about time-to-contact. In particular, the mean CEs in the dark are much smaller than use of tau alone would predict. Thus subjects do not rely on tau alone even when no other visual information about the situation is available.

The results obtained are thus in agreement with the following account of time-to-contact perception. In perceiving the time-to-contact (with an interception point) of a moving target which passes the observer some distance from the eyes,

subjects make use of the rate of change of direction of the target. Information about the rate of change of direction can be obtained either visually or through articular proprioception though visual information is more accurate and less subject to noise: the presence of visual information results in correspondingly more accurate and less variable performance in interceptive timing tasks.

It might be predicted that the speed of the ball would be correlated with the timing error for two reasons: (i) pursuit of a moving target is less accurate at higher target speeds, hence less well foveated. Note that pursuit eye movements lag consistently behind a visual target at speeds of 30 degrees of visual angle per second (Howard, 1982). In experiment 2a the ball moved such that its angular velocity at the eye was greater than 30 degrees/s. (ii) The faster the ball is going the less the time for which it is visible before it reaches the contact point (the latency of pursuit eye movements is about 125ms and it takes a further 100ms or so before the target is foveated, Howard, 1982). Indeed, two subjects volunteered the information that they sometimes found it difficult to visually locate the ball (foveate it); several other subjects when questioned about this confirmed that they too experienced similar difficulties. However, no correlations between ball speed and timing error were found for any of the subjects (see table 7.3). Nevertheless, the two subjects' reports that sometimes location of the target was difficult suggests the possibility that the effect of the peripheral lights may have been to improve visual tracking of the ball and as a consequence, the pick-up of time-to-contact information. If this were the case it might be expected that there would be a negative correlation between ball speed and timing error when only the ball was visible but no correlation when the peripheral lights were present in seating position B. Such a correlation over all subjects was insignificant ($r = -0.055$; $r^2 = 0.003$). However, this result does not rule out the possibility that the effect of peripheral lights was due to improved latency of target location and its subsequent pursuit. The next experiment is an attempt to determine whether this does account for the peripheral lights effect.

Experiment 2b

The difficulty in visually locating the ball reported by subjects in experiment 2a may have been due to the experimental set-up. The beginning of the ball's trajectory was obscured from view — the ball first appeared moving with a mean speed of about 1.1ms^{-1} when released 20cm behind the screen or about 1.4ms^{-1} when released 30cm behind the screen (these speeds are estimates calculated according to the model in figure 7.11). If the ball were visible from the beginning of its trajectory the difficulty in visually locating the ball and the delays due to pursuit latency and catch-up time might be effectively removed. In what follows a short experiment to test this possibility is reported. In two of the conditions in this experiment the whole of the ball's trajectory was visible to the subjects, in other respects the task was identical to that used in experiment 2a. If vision of the fixed surroundings provides information important for the control of interceptive timing then similar results to experiment 2a are expected. Failure to replicate the results might be due to vision of the initial position of the ball giving a cue to the time it will take to reach the interception point (this, it will be remembered, was one reason for obscuring the initial ball position from view in experiment 2a). However, since there was a measurable variability in the time taken to reach the interception point from any given starting point, use of this cue will have a certain error associated with it. The timing error should, therefore, be correlated with the ball speed if initial position is being used as a cue and no significant effects of peripheral environmental information are found (note that in experiment 2a essentially no correlation was found between timing error and ball speed).

Methods

Subjects Twelve subjects (all students and staff at the University of Edinburgh) participated voluntarily in this experiment. Six were male and six female. The age range was 21-28 years (to the nearest year) and the median age 22.5 years. All had normal vision or corrected to normal vision. None was informed of the hypothesis being tested and all stated that their preferred hand for catching and hitting was the right. Four subjects had previously participated in experiment 2a but had remained naive as to the hypothesis under study.

Apparatus The same apparatus that was used in experiment 2a was used.

Procedure Subjects sat in position B (figure 7.5) and performed the same trapping task as described above under three different conditions: 1) ball only visible from point of release (CD). 2) Ball visible from point of release and peripheral LEDs visible (CP). 3) Release point obscured as in experiment 2a and peripheral LEDs visible (OP). Balls were rolled from one of two initial positions on the trackway 10cms apart.

Subjects who had not participated in experiment 2a received a training session on the day preceding that on which the experiment was run which had the same structure as the training session given to subjects in experiment 2a. Subjects who had participated in experiment 2a did not receive this training session. All subjects received a practice session immediately before the experiment. Ten practice trials in full lighting conditions were followed by six trials in each of the experimental conditions presented in pseudorandom order. After a short break the experimental trials were run (subjects had been in the dark for about fifteen minutes by this time) and the intensity of the LEDs were adjusted in the same manner as in experiment 2a.

Experimental trials were run according to a counterbalanced order the same in principle to that used in experiment 2a. A total of fourteen trials were run in each condition. The experimental session was divided into two groups of 21 trials in a similar way to experiment 2a (seven trials in each of the three conditions in each of the two groups). Three different orderings of conditions were adopted to counterbalance for order effects, these were 1) CP, OP, CD; 2) CD, CP, OP; 3) OP, CD, CP. Subjects received seven trials in each condition in one of these orders (each order was given to four subjects), followed by a short break of about three minutes then seven trials in each condition in the reverse order. Errors were measured in the same way as in experiment 2a.

Results

The mean constant errors and variable errors are shown in the bar charts below (figures 7.14 and 7.15). In both figures it is apparent that the mean error when the trajectory was unoccluded and the peripheral LEDs were lit is less than the error when the peripheral LEDs were not lit. However, when part of the ball's trajectory was occluded both the mean CE and the mean VE are little different from the errors obtained when the trajectory was unoccluded and the LEDs were

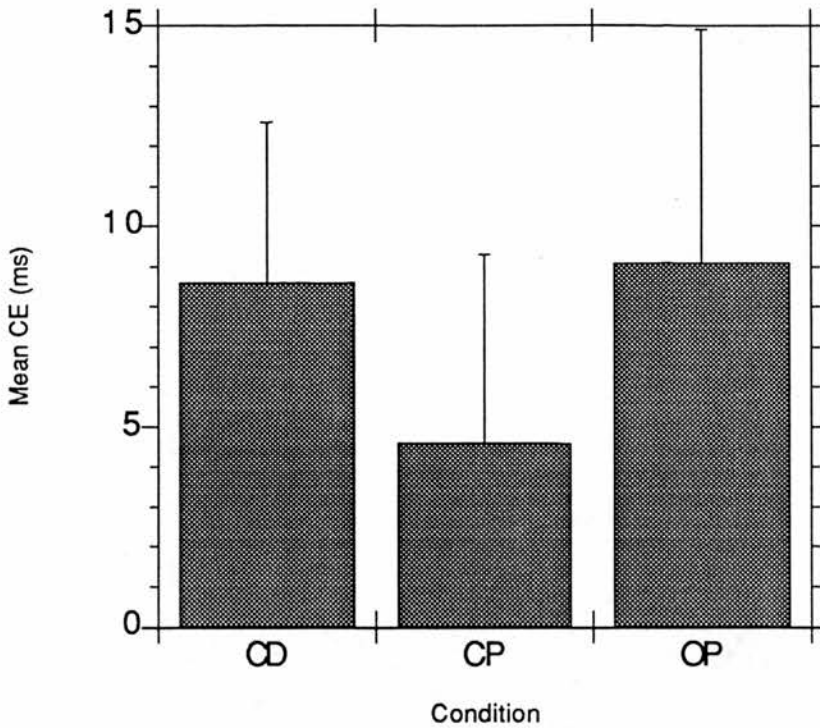


Figure 7.14. Mean constant error magnitudes (error bars show the standard deviations). Key: CD = trajectory unoccluded, only ball visible. CP = trajectory unoccluded, LEDs lit. OP = first part of trajectory occluded, LEDs lit.

unlit (the errors in the former condition are a little larger than in the latter).

Neither the CE nor the VE data was found to satisfy the assumptions required by parametric statistical tests. For this reason nonparametric statistical tests were used to analyze the data. The appropriate non-parametric test is the Friedman test (the non-parametric equivalent of a one-way repeated measures ANOVA). No significant main effect of viewing condition was found either for the CEs ($\chi_r^2 = 2.6$, $p > 0.1$, ns.) or the VEs ($\chi_r^2 = 4.6$, $0.1 > p > 0.05$, ns.). However, even though no overall main effect was found, it is possible that both types of error in the CP condition are significantly smaller than the errors in the CD condition (this is the effect that was being looked for). The significance of the difference in these errors was assessed using the Wilcoxon signed ranks test. A significant difference was found both for the CEs ($T = 17$, $p < 0.05$, one tailed) and the VEs ($T = 10$, $p < 0.025$, one tailed). The insignificance of the Friedman test is largely due to the very similar errors obtained in the CD and OP conditions

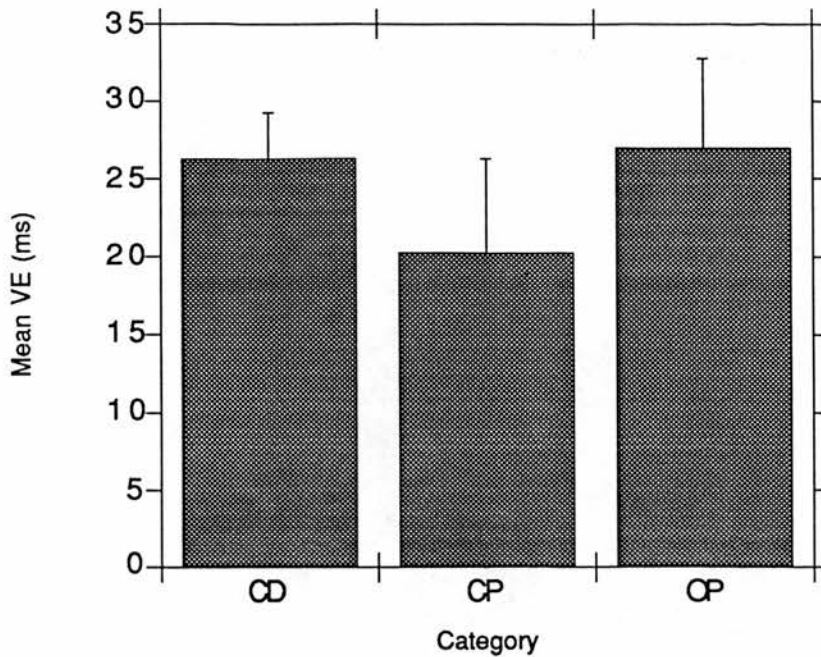


Figure 7.15. Mean variable errors (error bars show standard deviations). Key as for figure 7.14.

Discussion

The results show that the effect of minimal peripheral visual information about the fixed environment on timing error is preserved when no part of the moving target's path is occluded. This both replicates the finding of experiment 2a and indicates that the effect of the peripheral information is not to help the eyes locate and lock on to a moving target. The failure to find any difference between the condition in which only the ball is visible (CD) and the CP condition indicates that seeing the target from the beginning of its trajectory does improve performance. This could be due to one or more of the following: a) the initial position of the ball giving a small cue to time-to-contact; b) reduced latency of locating the moving target; c) longer viewing time. The results are therefore consistent with subjects reports in experiment 2a that balls were sometimes difficult to lock onto visually when the early part of the their trajectories were occluded. However, the results to tend to refute the conjecture that the effect of the peripheral visual information is due to it facilitating visual location of the moving target.

Experiment 3

The results of experiments 2a and 2b are consistent with the interpretation that a target's rate of change of direction is used in the perception of time-to-contact lending some support to the theory of what time-to-contact information might be used to time interceptions of objects which bypass an observer developed in earlier chapters. However, there is a possible alternative interpretation which prevents this conclusion being drawn unequivocally. It could be that the presence of the peripheral lights improves the subjects ability to pick-up information from the ball by improving tracking by pursuit eye movements. Such an interpretation is plausible because of an alternative source of time-to-nearest approach information described in von Hofsten and Lee (1985). Von Hofsten and Lee show that time-to-nearest approach, $T_n(t)$, is given by the relation,

$$T_n(t) = \frac{\tau_L}{2 + \dot{\tau}_L}. \quad (7.5)$$

This timing information involves only local tau and its time derivative, $\dot{\tau}_L$. If the subjects were making use of this information to time their interceptions, an interpretation of the peripheral light effect could be that the lights facilitate more precise pursuit of the ball which enables better detection of local tau and its time derivative. The following experiment was designed to test this possibility. The idea was to try and prevent pursuit eye movements playing a functional role in a timing task by asking subjects to stare straight ahead and not to follow the moving target with their eyes.

Methods

Subjects Twelve subjects (undergraduate and graduate students and research staff at the university of Edinburgh) participated voluntarily. Seven were male and five female. The age range was 21-36 years (to the nearest year) and the median age was 24.5 years. All had normal vision or corrected to normal vision. None was informed of the hypothesis being tested and all stated that their preferred hand for catching and hitting was the right. Three subjects had already participated in experiment 2b but remained naive as to the hypothesis under study.

Apparatus A new apparatus was constructed for this experiment and is shown in figure 7.16a. The illustrated apparatus was placed on two tables over which

black cloth was laid (black cloth also covered the wooden base of the apparatus). Black cloth was also hung around the apparatus; the complete set-up is shown schematically in figure 16b.

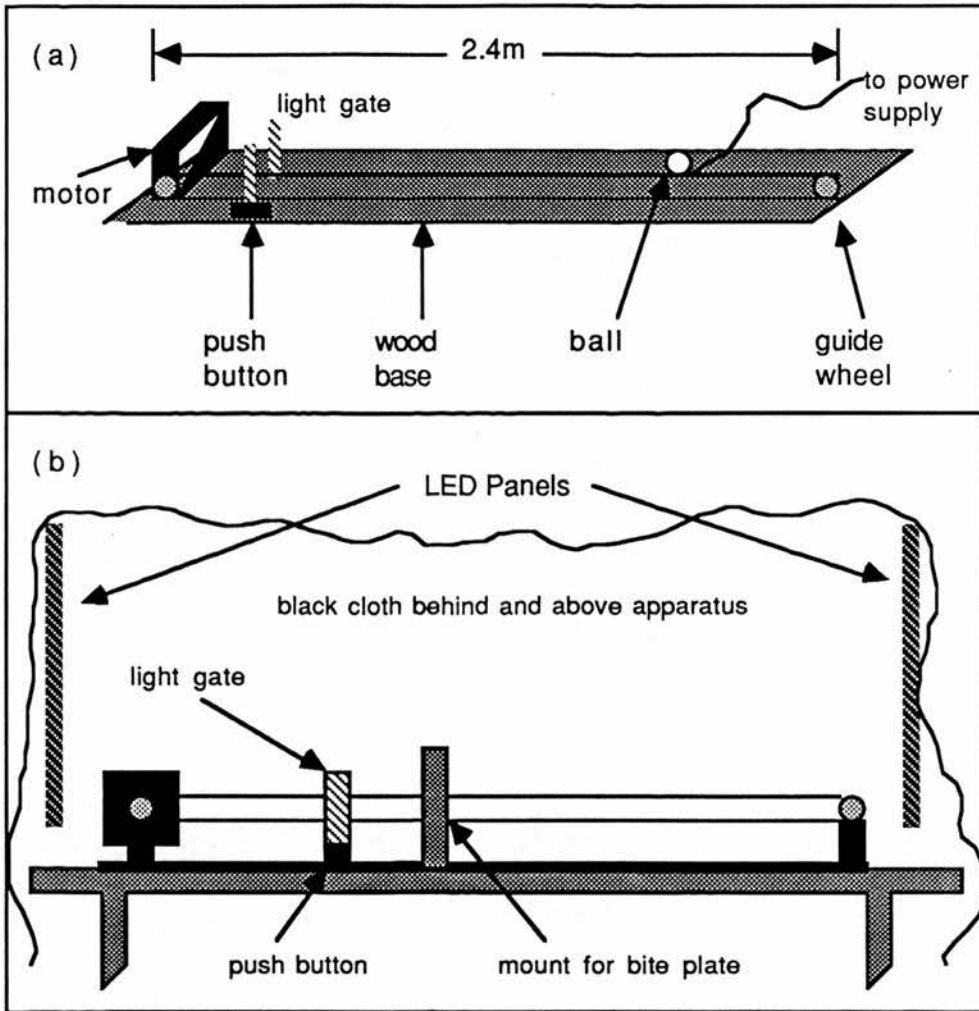


Figure 7.16. a) Apparatus for moving a luminous ball. A table tennis ball was threaded onto a wire running round two guide wheels, one of which was connected to a variable speed electric motor. Inserted into the ball through a small hole was an LED like those used in experiments 1 and 2 which was connected to a variable power supply by fine insulated wire. b) Experimental set-up: the same LED panels as used in experiment 2 were used here and held about 10cms above the table about 50cms from the ball moving apparatus on either side. The light gate was about 40cms from the motor and 30cms from the bite plate mount.

The time difference between the moment the ball passed through the light gate and the moment the button was pressed was recorded on an electronic millisecond timer with an addition which allowed the experimenter to determine whether the button had been pressed before or after the ball had passed through the gate.

Procedure Subjects sat facing the apparatus as illustrated in figure 7.16. Their heads were stabilized by requiring them to bite on a plate clamped to the mount illustrated and their eyes were about 35cms from the line of the moving target in this position. The apparatus illustrated in figure 7.16 was set up in a blacked-out room. Subjects held their right hands over the push button and their task was to slap their palm down on the button at the same moment as the moving target passed. The magnitude of the timing error was recorded directly on the millisecond timer and the sign of the timing error was indicated by an LED. Pilot study determined that people could perform sensibly on this task with timing errors in fully illuminated conditions being reliably within about 30ms.

In the experimental trials subjects were required to perform the task in the dark with the ball illuminated by the internal LED both with and without the peripheral panel LEDs illuminated keeping their eyes staring straight ahead (above the line of the moving target's path) with the instruction either to follow the target with their eyes or not to look at the target directly and to avoid following it with their eyes. In pilot study, subjects reported that they were able to avoid looking directly at the target and following it, though they were not convinced that they had succeeded in holding their eyes very steady. To assess the reliability of these claims horizontal eye movements of two pilot subjects were continuously monitored during performance of the task using an electro-oculographic (EOG) system the output of which could be viewed continuously on an oscilloscope screen allowing the experimenter to observe (and record) the subjects' eye-movements during performance².

Subjects performed six trials at each of four different target speeds (see below) presented in pseudorandom order. No signs of target directed horizontal eye movements were visually detectable from an oscilloscope display of the EOG output when the subjects were asked not to follow the moving target. During pursuit of the target, however, movements of the eyes were clearly visible on the display. Due to constraints on space and apparatus it was not possible to monitor subjects eye movements during the experimental trials. However, subjects were asked to verbally report any trials during which they found themselves looking at the target directly or following it with the intention of rejecting such trials as void and

² It was not possible to monitor vertical eye movements using the EOG equipment but these were of less concern.

rerunning them. In the event, no such reports were made by any of the subjects. The two subjects whose eye movements had been monitored in the pilot study also served as experimental subjects. Nine of the other ten subjects were tested with the EOG equipment (in the same manner as the pilot subjects described above) sometime after they had completed the experiment to see whether, as they had reported, they were able to avoid following the target during the experimental task. No subjects were observed to follow the moving target.

The experiment was organized according to a 2×2 factorial design with an eye movement factor (eyes following the target or eyes looking straight ahead) and an illumination factor (peripheral LEDs present or not present). There were thus four experimental conditions with a total of 24 trials in each condition. Four different ball speeds were used to add variability to the time taken for the ball to reach the interception point from the starting point. In each condition there were six trials at each speed. The order of speeds in each condition was determined by writing the four speeds on six pieces of paper each (giving 24 pieces of paper, one for each trial) and then drawing them blind from a container with the added constraint that no speed could be presented on more than two consecutive trials. The speed of the ball was measured using a pair of light gates separated by a known distance (cf. experiment 2). The four speeds measured using this technique were, approximately (to the nearest 5cms/s) 240, 160, 130 and 110cms/s. Each speed was associated with a particular setting of the dial of the potentiometer which controlled the motor speed.

Subjects who had not previously taken part in experiment 2 were given two practice sessions, the first one day prior to the experiment, the second immediately before the experiment. Subjects who had taken part in experiment 2 received only the practice session immediately before the experiment. In the practice session subjects received a total of sixteen trials³ in each of the experimental conditions (though they were not required to bite on the plate, they simply rested their chins on the bite plate mount). The fifteen trials were run consecutively and the order of conditions was different for each subject. Feedback about performance was given in the form of verbal information about whether the button had been pressed to early or too late and by how much. It was intended to reject any subject who failed

³ Four trials at each of four different speeds — the speeds used in practice were slightly different from those used in the experimental trials.

to perform sensibly on the task or who was unable to keep his or her eyes from following the target. In the event no one was rejected.

Experimental trials were grouped as follows. Six subjects received all the conditions in which the eyes were "fixed" followed by those in which the target was pursued, the other six received the reverse order. A short break of about three minutes separated the two sets of trials. Within each of the eye movement conditions the 24 trials in each illumination condition were grouped into three sets of eight trials run either in the order P, NP, P, NP, P, NP or its reverse. In each case three subjects in each group of six would get one of these orders and the other six the reverse order. This grouping of trials was adopted to control for order effects.

Results

As in experiments 2a and b, the timing error data was summarised as a constant error (CE) and a variable error (VE) for each subject. Descriptive statistics for the two measures are given in tables 7.4 and 7.5.

| Subject | lights+moving | lights+fixed | no lights+moving | no lights+fixed |
|----------|---------------|--------------|------------------|-----------------|
| mean | 20.2 | 29.1 | 18.8 | 37.3 |
| sd | 13.5 | 24 | 13.6 | 23.8 |
| skewness | 0.9 | 0.6 | 0.3 | 0.5 |
| kurtosis | -0.3 | -0.9 | -1.1 | -1.1 |

Table 7.4. Descriptive statistics for the unsigned constant errors (in milliseconds).

| | lights+moving | lights+fixed | no lights+moving | no lights+fixed |
|----------|---------------|--------------|------------------|-----------------|
| mean | 34.1 | 47.5 | 42.4 | 57.9 |
| sd | 8.6 | 11.9 | 8.6 | 12.8 |
| skewness | 0.2 | -0.1 | 0.1 | -0.6 |
| kurtosis | -1.4 | -0.8 | -0.6 | -0.3 |

Table 7.5. Descriptive statistics for the variable errors (in milliseconds).

The VE data indicate that errors are greater when the eyes are not following the target than when they are and are smaller when the peripheral LEDs are lit. This is shown graphically in figure 7.17. There is no apparent interaction between the illumination factor (peripheral lights lit or unlit) and the eye movement factor (eyes following target or not) and this is shown clearly in figure 7.18.

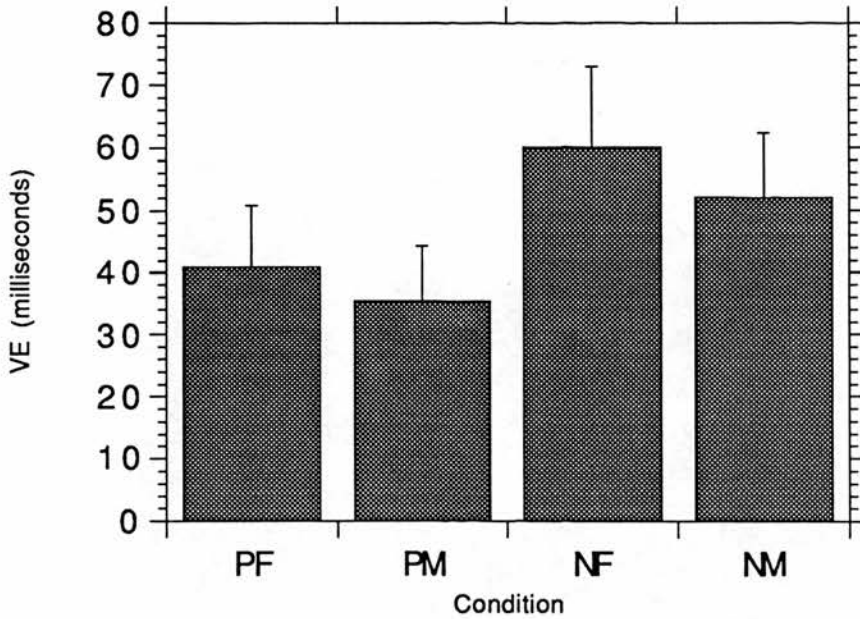


Figure 7.17. Bar chart of the mean VEs in the four conditions (error bars show the standard deviations). Key: PF = peripheral LEDs lit, eyes not following target. PM = LEDs lit, eyes following target. NF = LEDs not lit, eyes not following. NM = LEDs not lit, eyes following target.

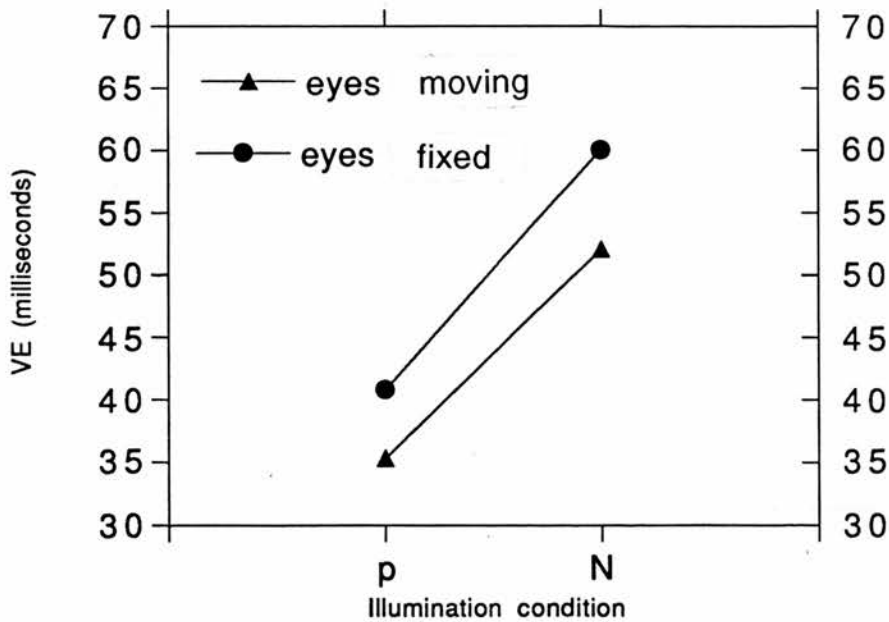


Figure 7.18. Graph showing the absence of an interaction between the illumination and eye movement conditions for the mean VEs. Key: P = LEDs lit, N = LEDs not lit.

The standard deviations of the VE data show a tendency to be proportional to the means, violating the homogeneity of variance assumptions underlying the

analysis of variance. This effect was found to be effectively removed by taking the natural logarithms of the original data (see table 7.6) and this transformed data was analyzed using a two-way repeated measures ANOVA. Both main effects were found to be statistically significant — errors were smaller when the eyes were moving ($F_{(1,11)} = 8.44, p < 0.05$) and when the peripheral LEDs were lit ($F_{(1,11)} = 42.51, p < 0.0005$). The interaction (see figure 7.18) was not significant ($F_{(1,11)} = 3.96, p > 0.05$). It was of interest to discover whether the effect of the peripheral lights was significant both with the eyes fixed and with them moving. A planned comparison simple main effect analysis was used to evaluate this. It was found that the peripheral LEDs had a significant effect both when the eyes followed the target ($F_{(1,11)} = 49.10, p < 0.0005$) and when they did not ($F_{(1,11)} = 42.19, p < 0.0005$).

| | lights+moving | lights+fixed | no lights+moving | no lights+fixed |
|------|---------------|--------------|------------------|-----------------|
| mean | 3.5 | 3.7 | 3.8 | 4.03 |
| sd | 0.25 | 0.21 | 0.27 | 0.25 |

Table 7.6. Means and standard deviations of log transformed variable error data.

The mean unsigned constant errors given in table 7.4 are plotted in figure 7.19. The errors are larger when the eyes are “fixed” than they are when the eyes are following the target. In the eyes fixed condition the errors are larger when the peripheral LEDs are unlit. Thus far the results are similar to experiment 2. When the eyes followed the target, however, there was very little difference between the means obtained under the two illumination conditions. As is clear from table 7.4, the constant error data departs from the conditions necessary for an analysis using parametric statistical tests. No simple transform was found that could effectively remove these departures from the parametric assumptions. Thus, as in experiment 2, non-parametric statistics were used to analyze these data. A Wilcoxon signed ranks test showed that there was no significant effect of the peripheral LEDs when the eyes were following the target ($T = 32$ ns.) but that there was a significant effect when the eyes were not following ($T = 10, p < 0.025$, one tailed).

Discussion

It was found that the presence of LEDs in the periphery of the visual field reduced both the variable timing error and the constant timing error when the subjects' eyes were not following the target, thus confirming the hypothesis that the effect

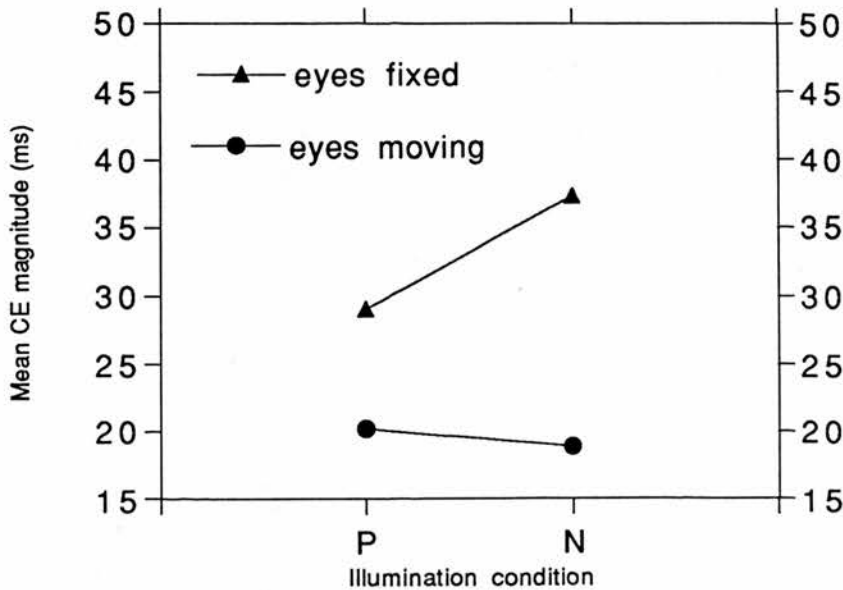


Figure 7.19 Plot of the mean unsigned constant errors. Key as figure 7.17.

of the peripheral lights is to inform about the rate of change of direction of a moving target. When the eyes were moving the presence of the peripheral LEDs was found to significantly reduce the magnitude of the variable error replicating the results of experiments 2a and b, but there was found to be no effect on the magnitude of the constant error. This result could be due to the difference in the task: both the constant errors and the variable errors tended to be larger in magnitude in the task used in this experiment than those obtained in the task used in experiments 2a and b. In both experiment 2a and experiment 2b the means of the constant error magnitudes were all less than 10ms whereas in this experiment these means were all well in excess of 10ms (table 7.4). A similar difference in the size of the variable errors exists. These differences are consistent with data from elsewhere: both McLeod, McLaughlin and Nimmo-Smith (1985) and Bootsma (1989) reported results showing that errors in a pseudo-interception task requiring a button press were greater than the errors obtained when the task involved actually intercepting a moving target. The effect of the button pressing task used here may have been, therefore, to obscure differences in timing errors. Thus, the failure to find a significant effect of peripheral LEDs on constant timing error in the eyes following target condition should not be interpreted as indicating

that such an effect does not exist in this or any other interception task. The results of this experiment are, therefore, in agreement with the results of the first three experiments and indicate that the peripheral LEDs provide subjects with information about a moving target's rate of change of direction which is important in timing interceptions. They contradict the alternative explanation that peripheral LEDs improve timing performance by aiding target foveation and visual pursuit.

Note also that the results are consistent with the idea that subjects are measuring the rate of change of direction of a moving target directly from retinal variables. A possible interpretation of the earlier experiments is that the rate at which peripheral visual texture comes into view or is occluded as the head turns provides the subject with information about the rate of head rotation in the environment and this is combined with the rate of eye rotation in the head to compute the rate of change of target direction. Since the head was completely stationary in this experiment, this cannot be the explanation of the results reported here.

§7.3. INVESTIGATION OF THE EFFECT OF VISION OF THE INTERCEPTION POINT ON INTERCEPTIVE TIMING

The results of the experiments reported in the previous section support the contention that the rate of change of direction of a moving target is important in the control of interceptive timing. This in turn is consistent with the use of the information about time-to-contact described in chapter 4 which requires perception of the rate of change of target direction. Two types of time-to-contact information involve this variable, the time-to-nearest approach information introduced by Lee and Young (1985) and the general time-to-contact information introduced in chapter 4. The experiments results so far reported are consistent with the use of either of these two sources of information. In this section the results of two experiments are reported which attempt to determine whether the perception of the interception point is important in the control of interceptive timing as the timing information introduced in chapter 4 would predict.

Experiment 4

Lee (Lee & Young, 1985; von Hofsten & Lee, 1985) suggested that if tau alone were not sufficient to account for the skilled timing of interceptive acts like catching then the required information might be provided by time-to-nearest-approach (see equations 4.21 and 7.5 above) which does not require that the point of interception be perceived. If this is the information used by human subjects to time interceptions of moving targets which pass the subject some distance from the eye, then interceptions of such targets not made at the point of nearest approach will be subject to errors in timing.

Consider the trapping task used in experiment 2a and the model of it described in the discussion of that experiment (see figure 7.11). What error would be expected if the trap were required in front of or behind the point of nearest approach on the assumption that a person will use time-to-nearest approach information to time the trapping movement? Considering the model in figure 7.11, suppose the trap is to be made a distance d in front of the point of nearest approach.

The ball starts from rest a distance 2.25 metres from the point of nearest approach and hence the time to nearest approach (T_n) may be obtained from the equation

$$2.25 = \frac{1}{2}aT_n^2.$$

The time-to-contact (T_c) with the actual trapping point may be obtained from the equation

$$2.25 - d = \frac{1}{2}aT_c^2.$$

As in experiment 2a, let the acceleration a be equal to $3.0ms^{-2}$. The timing error e resulting from using T_n as an estimate of T_c is defined to be $T_c - T_n$. Its value depends on d and is given by

$$1.22 - \sqrt{(1.5 - 2d/3)} \tag{7.6}$$

which follows immediately from the two equations for T_n and T_c . If d is 0.4 metres the timing error is about 110 milliseconds.

If the subject uses the constant velocity strategy, errors tend to be larger. The actual time-to-contact with the interception point (parameters as above with $d = 40\text{cms}$) and the time-to-nearest approach assuming constant velocity are plotted in figure 7.20, along with their difference (the error function), as a function of distance from the point of nearest approach. The error is at a minimum when the target is at the interception point — about 116ms.

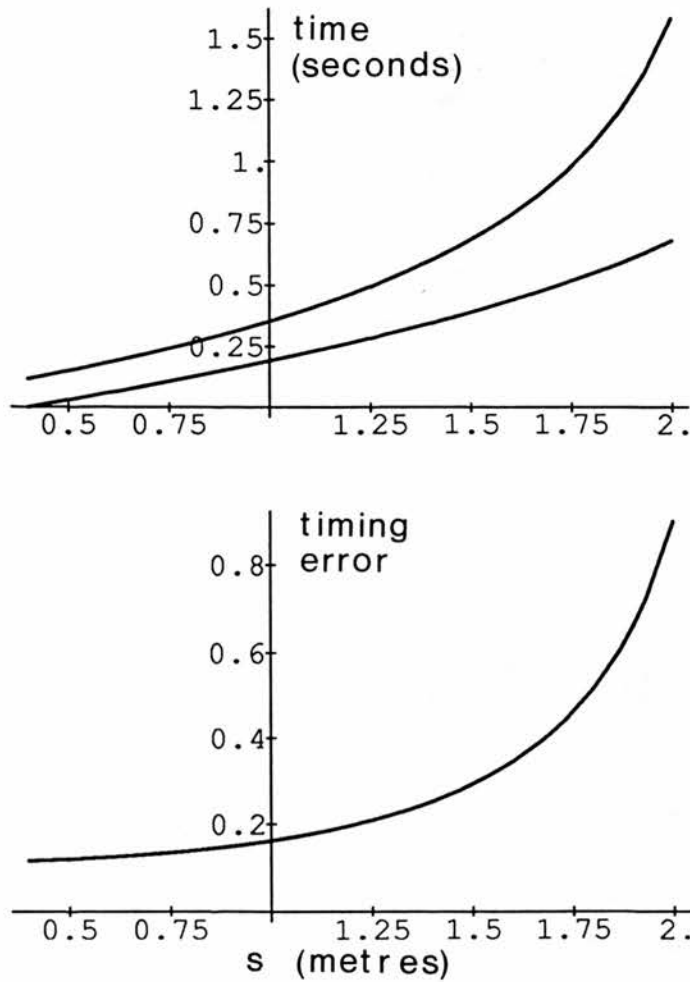


Figure 7.20 a) Lower curve shows the actual time-to-contact (T_c) with an interception point 40cms in front of the point of the point of nearest approach plotted as a function of distance (s) from the point of nearest approach, the upper curve shows the time-to-nearest approach assuming the constant velocity strategy. b) The difference between the upper and lower curves in a).

The following experiment set out to investigate whether time-to-nearest approach information is sufficient to account for performance in a task similar to those used in the experiments reported above. Timing performance was assessed at three different points using an apparatus similar to that used in experiment 2: the point of nearest approach, a point 40cms in front of this point and a point 40ms behind it.

Methods

Subjects Ten subjects (undergraduate and graduate students and research staff at

the university of Edinburgh) participated voluntarily. Seven were male and three female. The age range was 21-45 years (to the nearest year) and the median age was 24 years. All had normal vision or corrected to normal vision. None was informed of the hypothesis being tested and all stated that their preferred hand for catching and hitting was the right. All subjects had already participated in one or other of the previous experiments but were naive as to the hypothesis under study here.

Apparatus The apparatus used in experiments 2a and b was modified for used in this experiment. The perspex sheet was fixed in position over the trackway and could not be pressed down so as to trap the ball. A single moveable light-gate was made that could be placed in any position over the trackway under the perspex sheet. The push button used in experiment 3 was placed over the trackway on top of the perspex. With the button placed perpendicularly above the light gate, the time between when a ball passed through the gate and a button press could be measured directly as in experiment 3.

Procedure Subjects sat in position B marked in figure 7.5 (experiment 2a). The subjects' heads were held in a helmet attached to a fixed beam to ensure that the eyes did not move any appreciable distance in space during the trials. The centre of the helmet was about 55cms from the trackway. This procedure was used so as to maintain a more or less constant point of nearest approach of the moving target to the subjects' eyes over trials. The subjects held their hands just over the push button and their task was to bring their hands down on the push button as if they were trying to trap the ball. The task was performed in the dark with self-luminous balls as in the previous experiments. Subjects were required to perform the task with the push button in three different positions — at a position within 5cms of the point of nearest approach to the eyes (position 2), at a position 40cms behind the point of nearest approach (position 1) and at a position 40cms in front of the point of nearest approach (position 3). In each position subjects performed the task either with a lit LED on the back of their responding hand (condition H) or with the LED somewhere on the beam which supported the perspex sheet in experiments 2a and b (see figure 7.5) (condition N). The position of this LED was varied from trial to trial so that it did not bear a fixed spatial relationship to the three "interception" positions.* The purpose of the LED on the hand was to

*It was, in fact, moved to a new position on each trial.

determine whether vision of the hand aids timing in this task but the possibility that the LED could also be informing about the rate of eye rotation needed to be controlled for, hence the presence of the lit LED on the beam when the LED on the hand was unlit.

A total of 42 trials were run in each of three positions, 21 with the LED on the hand lit, 21 with the LED on the beam lit. The order of positions was determined by writing the numbers of each condition on 42 small pieces of paper and drawing them blind from a container for each subject with the added constraint that no position could appear more than twice in row. The 126 trials were divided into 18 sets of 7 each set being either LED on hand lit (H) or LED on beam lit (N). Nine sets were run followed by a short break of about five minutes followed by the remaining nine sets. Five subjects received the sets ordered N, H, N, H etc. and five subjects received them in the reverse order (H appearing first). All subjects received one practice session immediately before the experimental trials. Five practice trials were given in each of the experimental conditions, presented in a pseudorandom order determined by a draw.

Results

The mean constant errors for the individual subjects are presented in table 7.7a and the overall descriptive statistics of these data are presented in table 7.7b. Table 7.7a shows that eight of the ten subjects displayed a bias towards responding late (negative timing error) which held over all conditions and is reflected in the overall means (table 7.7b). Two subjects displayed a bias towards early responding (positive timing error) which also held over all conditions.

There was little difference in the constant errors between conditions. The overall mean of the CEs in the three no LED conditions was exactly the same as the overall mean in the LED conditions (-10.6 ms in each case). The overall mean of the magnitudes of the CEs in the no LED conditions was also the same as the overall mean of the magnitudes in the LED conditions (20ms in each case). The means in table 7.7b suggest that timing accuracy in position 1 might have been slightly superior to the accuracy in the other two positions. However, as can be seen from table 7.7a, this effect is entirely due to the large positive errors of subjects 3 and 4 in position 1 biasing the means. If we consider the means of the constant error magnitudes, almost the reverse is found (mean magnitudes in

a)

| Subject no. | N1 | N2 | N3 | H1 | H2 | H3 |
|-------------|-----|-----|-----|-----|-----|-----|
| 1 | -4 | -3 | -4 | -3 | 3 | -11 |
| 2 | -15 | -28 | -31 | -6 | -23 | -20 |
| 3 | 28 | 12 | 14 | 9 | 5 | 17 |
| 4 | 41 | 35 | 8 | 44 | 27 | 6 |
| 5 | -19 | -31 | -18 | -2 | 0 | 8 |
| 6 | 4 | -19 | -18 | -29 | -31 | -34 |
| 7 | -12 | -8 | -10 | 20 | -27 | -16 |
| 8 | -29 | -18 | -23 | -27 | -27 | -30 |
| 9 | -54 | -25 | -6 | -31 | -24 | -17 |
| 10 | -26 | -27 | -31 | -23 | -42 | -35 |

b)

| Condition | N1 | N2 | N3 | H1 | H2 | H3 |
|--------------|-------|-------|-------|-------|-------|-------|
| mean | -8.6 | -11.2 | -11.9 | -4.8 | -13.9 | -13.2 |
| sd | 27.7 | 22.8 | 15.3 | 24.3 | 21.4 | 18.2 |
| skewness | 0.4 | 1.2 | 0.3 | 0.7 | 0.6 | 0.4 |
| kurtosis | -0.4 | 0.3 | -1.0 | -0.4 | -0.7 | -1.1 |
| overall mean | -10.6 | | | -10.6 | | |

Table 7.7 a) Mean constant errors (in milliseconds) for the individual subjects. Key: N1 = no LED on hand, position 1; N2 = no LED, position 2; N3 = no LED, position 3; H1 = LED on hand, position 1; H2 = LED, position 2; H3 = LED, position 3. b) Descriptive statistics for the constant errors. Key as above.

milliseconds are as follows: N1 = 23, N2 = 21, N3 = 16.3, H1 = 19.4, H2 = 20.9, H3 = 19.4). There are no discriminable effects of hand position or LED position on timing accuracy measured as constant error. The errors are much smaller, however, and do not follow the pattern of sign change that they should show if subjects were using time-to-nearest approach to time their responses. This failure to follow the pattern predicted by the hypothesis that subjects use time-to-nearest approach information is shown dramatically in figure 7.21.

The descriptive statistics for the subjects variable errors are presented in table 7.8 and the means are presented as a bar chart in figure 7.22. As shown clearly in figure 7.22, the VEs are smaller when the LED was on the hand than when it was not in all three positions.

The standard deviations of the VE data display a tendency to be smaller when

| Condition | N1 | N2 | N3 | H1 | H2 | H3 |
|-----------|------|------|------|------|------|------|
| mean | 36.5 | 37.1 | 38.8 | 27.4 | 32.8 | 31.4 |
| sd | 11 | 9.7 | 9.4 | 6.6 | 9.6 | 5.1 |
| skewness | -0.1 | -0.6 | 0.7 | 0.3 | 0.2 | -1.1 |
| kurtosis | -0.8 | -0.3 | 0.5 | -1.4 | -1.1 | -0.2 |

Table 7.8. Descriptive statistics for the variable errors (milliseconds). Key as for table 7.7

the mean is smaller (table 7.8) thus violating the homogeneity of variance requirements of the analysis of variance. This tendency was found to be reduced by taking the natural logarithms of the individual subject VEs (table 7.9). This transformed data was analyzed using a 2×3 repeated measures ANOVA. The LED position main effect was statistically significant ($F_{(1,9)} = 12.66, p < 0.01$) but neither the interception position main effect ($F_{(2,18)} = 1.49, p > 0.05$) nor the interaction ($F_{(2,18)} = 2.8, p > 0.05$) were statistically significant.

| Condition | N1 | N2 | N3 | H1 | H2 | H3 |
|-----------|------|------|------|------|------|------|
| mean | 3.55 | 3.58 | 3.6 | 3.29 | 3.45 | 3.43 |
| sd | 0.34 | 0.31 | 0.24 | 0.24 | 0.31 | 0.18 |

Table 7.9 Means and standard deviations of the log transformed VE data. Key as for table 7.7

Discussion

The results show that subjects do not rely on time-to-nearest approach information to time interceptions at points other than the point of nearest approach. The timing errors were simply far too small and did not display an appropriate pattern — if time-to-nearest approach information was being used, then responses made

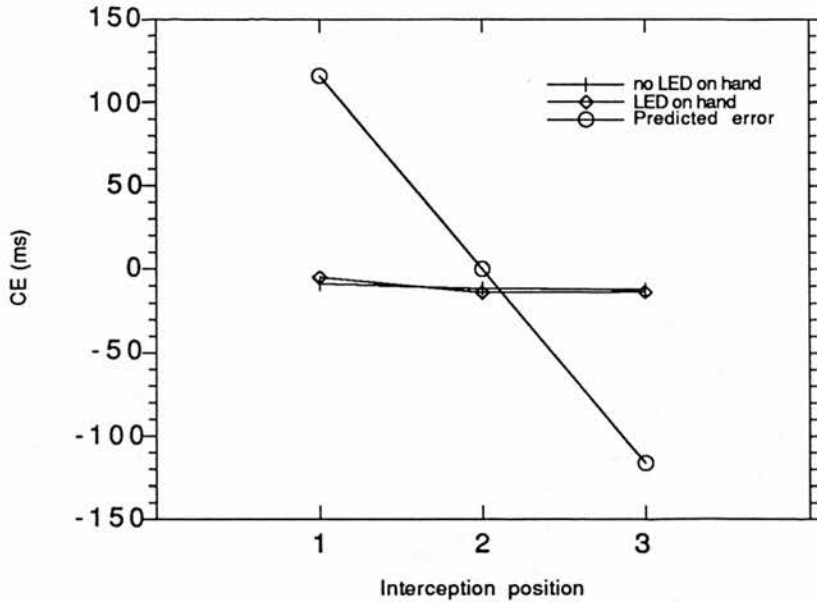


Figure 7.21. Plot showing how the CEs obtained compare with those predicted by the hypothesis that subjects use time-to-nearest approach information.

at points past that of nearest approach should be too early (positive timing error) and they should be too late (negative timing error) at points in front of the point of nearest approach. This pattern was not observed — there tended to be no difference in the sign of the constant errors between points in front of and behind the point of nearest approach.

When the subjects' hands were marked with an LED (effectively marking the pseudo-interception point) the variable error was reduced but no effect on the constant error was detected. This result indicates that vision of the interception point is of some use in interceptive timing. The results of this experiment are, therefore, inconsistent with the use of time-to-nearest approach information which predicts no effect of seeing the interception point. They are, however, consistent with the

use of the more general time-to-contact information described in chapter 4. It is difficult to conclude with certainty what the role of vision of the interception point is based on these results since it has only been shown to have the effect of slightly reducing the variable timing error. Use of the time-to-contact information described in chapter 4 implies that accurate timing depends upon accurate perception of the interception point and hence that misperception of the intercept should adversely affect the constant timing error. In the next experiment we set out to determine whether this is the case.

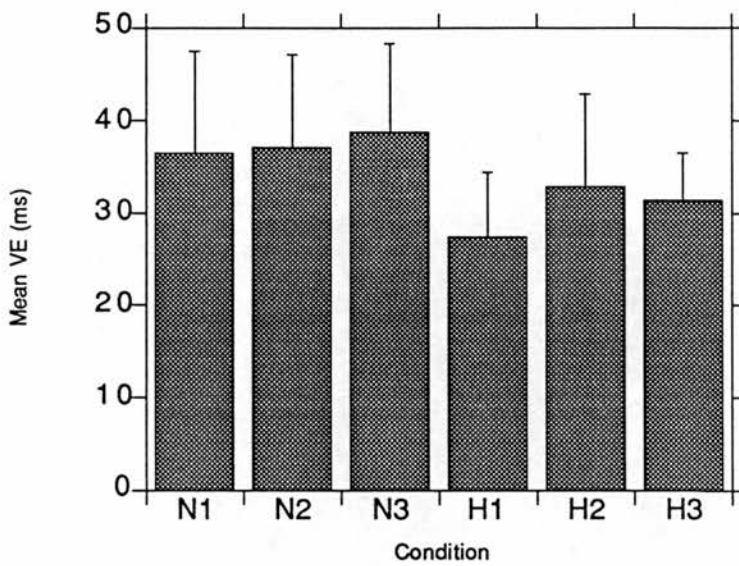


Figure 7.22 Bar chart of the mean variable errors (error bars show the standard deviations. Key as for table 7.6.

Experiment 5

Normally, the position of a part of the body perceived visually corresponds to its position perceived via articular proprioception. This correspondence is disrupted when a subject wears left-right displacing prisms. Such prisms allow for a direct investigation of whether perception of the interception point is important in the control of interceptive timing. Consider the following figure (figure 7.23).

If the observer is wearing left-right displacing prisms and can see both the moving object and the interception point then the prisms should have no effect on the observer's judgement of time-to-collision⁴. However, if the observer can only see the moving object and perceives the position of the interception point through articular proprioception (e.g., the observer's hand is placed at the interception point) then left-right displacing prisms will have an effect on time-to-collision judgement if the subject uses articular proprioceptive information about the interception point to estimate time-to-collision. The angle $\theta(t)$ between the moving object's position and the interception point will be increased if the prisms shift to the left and decreased if the prisms shift to the right in figure 7.23. A corresponding increase or decrease in the estimated time-to-collision is to be expected.

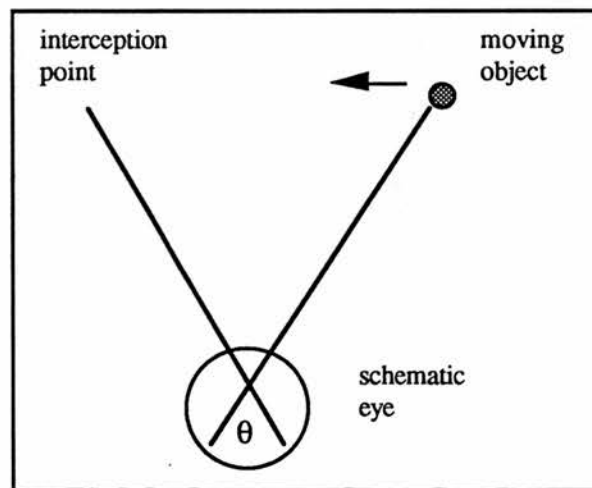


Figure 7.23. Visual target (moving object) moving towards a visually specified interception point.

⁴ There will only be no effect if the optical action of the prisms is simply to shift visual space to the left or right relative to articular proprioceptive space. Typically, however, prisms distort visual space in other ways which may interfere with time-to-contact judgements. Such effects may be assumed small compared with the shift effect of the prisms.

It is clear also that different target speeds should affect the timing error induced by the displacing prisms. To compute roughly how large such errors would be using a set-up like that described in experiment 3, consider the geometry of the situation illustrated in figure 7.24 (the parameters are roughly those that used in the experiment reported below). By application of the sine rule the point at which the target appears visually to be at the proprioceptively specified interception point is found to be displaced approximately 18.5 cms to the right of the interception point (assuming a 15° angle of displacement by the prism). If the target speeds are the same as those used in experiment 3 — 240, 160, 130 and 110 cms/s — then the corresponding timing errors induced by wearing prisms are expected to be 77, 115, 142 and 168 ms respectively. The timing error is inversely proportional to the speed of the target. These predictions were investigated in the experiment reported below.

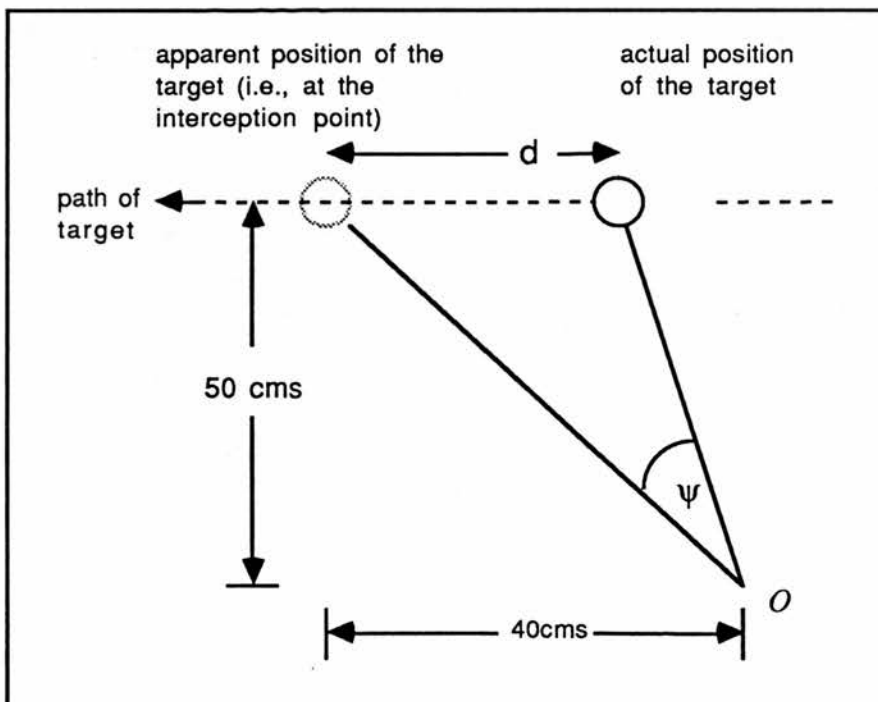


Figure 7.24. Geometry of interceptive timing task. The prism displacement angle $\psi = 15$ degrees and the distances shown are roughly what they were in experiment 3.

Methods

Subjects Twelve subjects (undergraduate and graduate students and research staff at the university of Edinburgh) participated voluntarily. Seven were male and five

female. The age range was 21-29 years (to the nearest year) and the median age was 22.5 years. All had normal vision or vision corrected to normal with contact lenses (none wore glasses). None were informed of the hypothesis being tested and all stated that their preferred hand for catching and hitting was the right. Six subjects had already participated in one or other of the earlier experiments but remained naive as to the hypothesis under study.

Apparatus The same apparatus as that used in experiment 3 was used, but without the bite plate and support.

Procedure Subjects sat in the same position as that described earlier in experiment 3. They were not, however, required to bite on a plate. Subjects performed monocularly the same timing task as that described in the report of experiment 3 above. Subjects' left eyes were occluded with an eye patch. The experiment was organized according to a 2×2 repeated measures factorial design. The two factors were prisms — subjects either wore a displacing prism over their unoccluded eye (P) or they did not (NP) — and interception point specification — an LED placed above the interception point as defined by the position of the light gate was either lit (L) or unlit (D).

Practice was organised similarly to previous experiments. If a subject had not participated in an earlier experiment, he or she was given two practice sessions, one on the day preceding the experiment and one immediately before it. Other subjects received only a practice session immediately before the experiment. Subjects did not wear prisms during any of the practice trials but they did wear the prism glasses without the prisms in. A practice session consisted of 30 trials with the LED lit (L) and thirty trials with it unlit (D); fifteen trials in each case were binocular and fifteen monocular. It was intended to reject any subject who failed to perform sensibly on the task. Again, in the event no one was rejected.

Experimental trials were divided into four groups of 48 trials each. In two of the groups subjects wore the prism glasses with the right eye prism in place (condition P), for the other two groups of trials the prism was removed (condition NP). Six subjects received the groups of trials in the order P, NP, P, NP; the other six received them in the reverse order. In each group of trials the LED over the "interception" point was lit (condition L) on twenty four trials and unlit (condition D) on the other twenty four. Each group of 48 trials was organised into four sets

of six in one or other of the two LED conditions; subjects received these sets in one of two orders — L, D, L, D or the reverse. A sequence of trials run according to this scheme might be as follows:

| | | | | |
|-----------------|---------|---------|---------|---------|
| Prism Condition | P | NP | P | NP |
| LED Condition | L D L D | D L D L | D L D L | L D L D |
| No. Trials | 6 6 6 6 | 6 6 6 6 | 6 6 6 6 | 6 6 6 6 |

The same four speeds as were used in experiment 3 were used here also. In each condition the subject would receive three trials at each speed presented in an order determined by a draw as in experiment 3. This scheme just described was adopted to counterbalance for order effects including any that might be the result of adaptation to the prism. Pilot study did indicate, however, that adaptation to prisms under the experimental conditions was negligible, as might be expected (the subject has no visual information about his or her body position which can be compared to articular proprioceptive information and drive adaptational processes).

Results

The mean constant timing errors and variable timing errors are tabulated in tables 7.10 and 7.11 and plotted in figures 7.25 and 7.26 respectively.

| | Prisms+light | Prisms+no light | No prisms+light | No prisms+no light |
|------|--------------|-----------------|-----------------|--------------------|
| mean | -36.700 | 57.500 | -28.170 | -11.500 |
| sd | 24.500 | 26.300 | 20.400 | 26.200 |

Table 7.10. Mean constant errors (in milliseconds)

| | Prisms+light | Prisms+no light | No prisms+light | No prisms+no light |
|------|--------------|-----------------|-----------------|--------------------|
| mean | 43.7 | 46.8 | 38.6 | 40.3 |
| sd | 4.4 | 7.0 | 7.1 | 6.1 |

Table 7.11 Mean variable errors (in milliseconds)

From the graphs it can be seen that wearing the prisms appears to have little effect on the variable error (a small increase in the means is apparent). Marking the interception point with an LED does not appear to affect the variable error

very much either, though a small decrease in the mean VE in the presence of the marker LED is apparent (figure 7.25). The picture for the constant errors is quite different. When the subject wore the prisms and the interception point was not visible the errors obtained are large and positive (button pressed too early) whereas they tended to be much smaller and negative (button pressed too late) when the interception point was visually marked by the LED. Without the prisms the mean CEs are negative whether the interception point was marked or not and the mean when the intercept was unmarked is smaller than that obtained when it was marked. In the latter case the mean CE is not much different from that obtained when the subjects wore prisms and the intercept was marked.

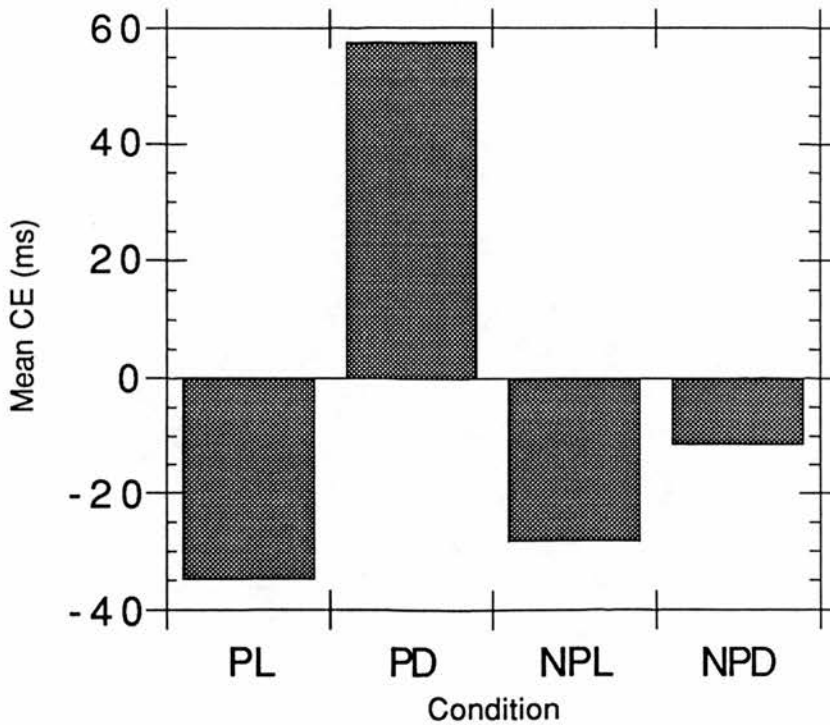


Figure 7.25. Bar chart of the mean constant errors in the four experimental conditions. Key: PL = prisms worn and LED marking interception point. PD = prisms worn, no LED. NPL = no prisms worn, LED. NPD = no prisms worn, no LED.

The mean VE data was analyzed using a 2×2 repeated measures ANOVA. It was found there was an effect due to wearing of the prisms ($F_{(1,11)} = 12.44$, $p < 0.01$) — VEs were larger when the subjects wore the prism. There was no significant effect of the marker LED ($F_{(1,11)} = 1.59$, $p > 0.1$) and no significant interaction between the two factors ($F_{(1,11)} = 0.12$, $p > 0.1$).

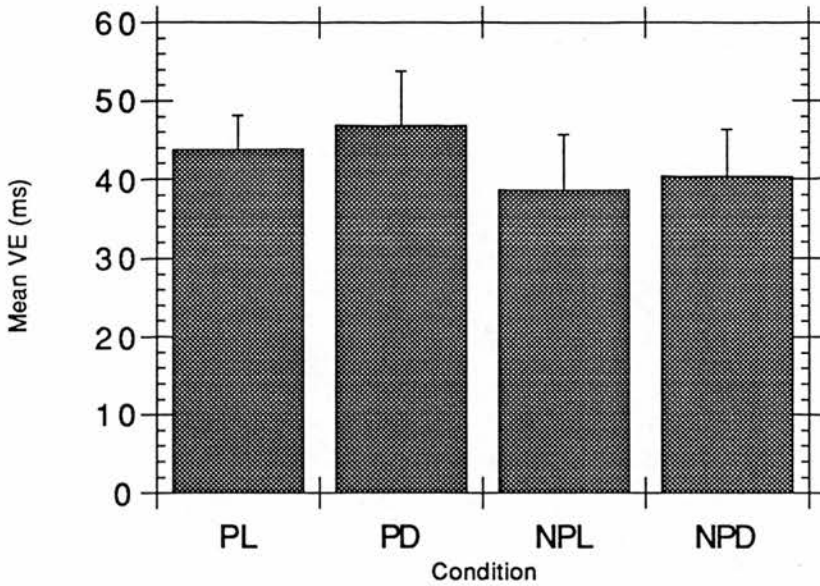


Figure 7.26. Bar chart of the mean variable errors in the four experimental conditions. Key as for figure 7.25.

| Subject no. | Speed 1 | Speed 2 | Speed 3 | Speed 4 |
|-------------|-------------|-------------|-------------|---------------|
| 1 | 98.5 | 96.0 | 76.5 | 48.0 |
| 2 | 107.0 | 76.0 | 53.5 | 58.5 |
| 3 | 100.5 | 74.0 | 92.5 | 54.0 |
| 4 | 61.5 | 51.0 | 27.5 | 5.5 |
| 5 | 98.0 | 92.0 | 92.5 | 20.5 |
| 6 | 48.0 | 69.5 | 10.0 | -4.5 |
| 7 | 60.0 | 13.0 | 24.5 | -21.5 |
| 8 | 69.5 | 31.5 | -9.5 | -18.5 |
| 9 | 91.5 | 59.5 | 34.5 | 12.5 |
| 10 | 106.5 | 67.5 | 71.5 | 22.0 |
| 11 | 89.5 | 96.5 | 94.5 | 72.5 |
| 12 | 120.5 | 107.5 | 79.5 | 37.0 |
| mean | 87.6 | 69.5 | 54.0 | 23.8 |
| sd | 22.5 | 27.8 | 35.7 | 30.700 |

Table 7.12. Median timing errors for the individual subjects at each of the four speeds in the prism condition. The data are summarized by the means and standard deviations.

The timing errors at each speed for each subject when the prisms were worn were investigated (six trials at each speed). The median timing error was chosen

as the best measure of central tendency for this data. The median timing errors for the individual subjects along with their means and standard deviations over all subjects are shown in table 7.12. The means are plotted as a function of target speed in figure 7.27, a least squares fit of the data to a power law model ($y = ax^b$) is shown. The equation of the curve is $y = 107x^{-1.67}$ and the value of r^2 for the fit is 0.98 (the fit accounts for 98% of the variance in the data).

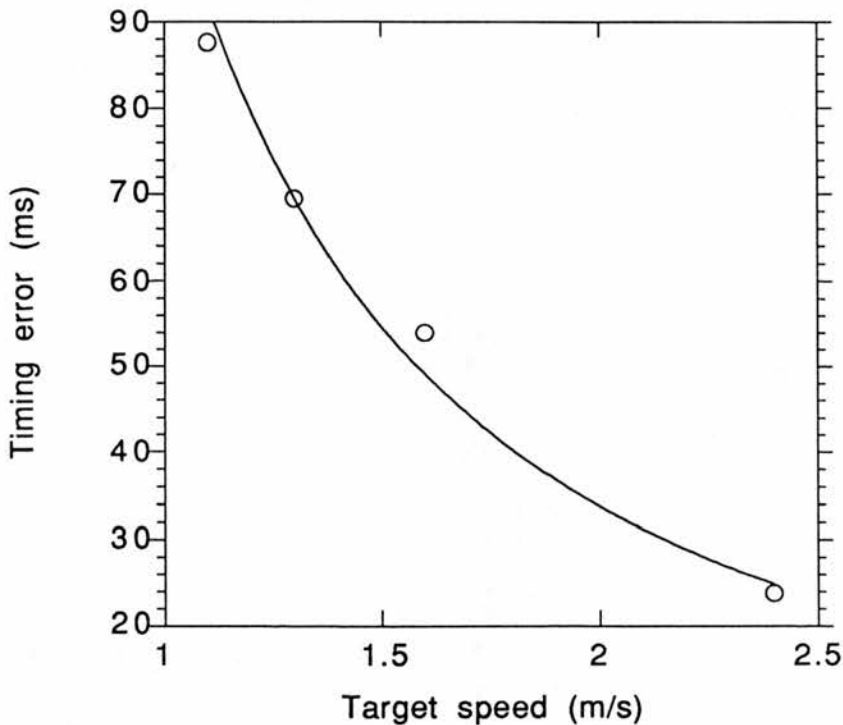


Figure 7.27. Means of the individual median timing errors in the prism condition plotted as a function of target speed. The curve shown is the least squares power law fit to the data.

Discussion

Subjects generally displayed a bias to respond too late (negative constant error) in the pseudo-interception task used in this experiment just as they did in experiments 3 and 4 where a similar task was used. However, when a prism was worn the constant errors changed to being large and positive. This result is exactly what would be predicted were subjects using the time-to-interception point information described in chapter 4 to time their responses. The timing errors decreased as a function of target speed when the subjects wore the prism as expected, but the exponent of the best fit power law of the form $y = ax^b$ was not -1 as predicted

but -1.67 . Considering the variability of the data (the variable errors are large and significantly larger when prisms are worn) this is in surprisingly good agreement with the theory. The actual mean timing errors for the four speeds in the prisms condition were smaller than those predicted by 50–60 ms in each case. This is probably due largely to the bias to respond late in this task which results in a negative error when the prisms are not worn and a smaller than expected error when the prisms are worn.

The results support the idea described in chapter 5 that subjects might be able to use more than one perceptual system to obtain time-to-contact information. The fact that the prisms had such an unequivocal effect on timing performance implies that not only is the interception point important in interceptive timing, but its position can be obtained using articular proprioception if vision is not available. In conclusion, the results of this experiment are in good qualitative agreement with the predictions based on the use of the interception point to time interceptions. They are, therefore, consistent with the use of the time-to-interception point information described in chapter 4 but not with the use of tau alone or with the use of the time-to-nearest approach information described by Lee and Young (1985).

§7.4. CONCLUSIONS

In this chapter the results of six experiments have been presented which support the hypothesis that both a moving target's rate of change of direction and position of the interception point are used in timing interceptions of targets which are not on a collision course with the observer's eyes. The results are therefore consistent with the use of the information described in chapter 4 equation 4.23 to obtain time-to-contact information and inconsistent with the use of local tau alone or with time-to-nearest approach information. The results do not firmly establish that it is the information described in chapter 4 that subjects are using; establishing this is actually extremely difficult and the problem is discussed further in the next chapter. It is, however, clear that neither local tau on its own nor time-to-nearest approach information is sufficient to account for the timing of interceptions in the tasks used in the experiments reported here.

The results also support the idea, described in chapter 5, that timing information can be obtained using more than one perceptual system. The results show that although timing performance is both

more accurate and more precise when visual information about the interception point and the rate of change of target direction is present, in the absence of visual information performance does not degrade to the level that would be expected were timing based only on the visual information remaining. This suggests that the rate of change of direction and interception point information can be obtained using articular proprioception as suggested in chapter 5. Experiment 5 illustrates this most convincingly: the timing accuracy of monocular subjects is severely affected by wearing displacing prisms when the interception point is invisible, and must therefore be perceived using articular proprioception, but relatively unaffected by the prisms when the interception point is visible. In conclusion, the results of the experiments reported here lend support to the theory, developed in chapters 4 and 5, of how time-to-interception point information might be obtained by human subjects. It would be desirable to obtain more quantitative results to test the theory more rigorously, but the tasks employed in the experiments reported here give rise to data that is probably too variable to use to test quantitative predictions. This problem is discussed further in the next chapter.

8

Conclusions

§8.1. OVERALL SUMMARY

The conclusions of each of the foregoing seven chapters may be summarized briefly as follows:

Chapter 1: the problem of how interceptive acts are perceptually timed has long been associated with the question of how best to characterize the nature of perceptual processes in animals and people. Much discussion of perceptual timing assumes that perception is best characterized as “direct”.

Consequently, many of the research questions concerning perceptual timing have centred on whether the human visual system is best described as using a direct method (based on tau) or some kind of indirect or computational method. In chapter 1 it was argued that the directness or indirectness of perception is not an empirically meaningful question and further, that it is founded on inappropriate notions of what constitutes computation.

It was concluded that the modern computational approach to perception is not distinguishable from direct perception when it comes to the question of what information is used to time interceptions and how that information is extracted by the perceptual systems. As a consequence it was argued that research into the obtaining of information for interceptive action should not be directed at trying to determine whether perception is direct or indirect.

Chapter 2: here a framework for posing questions concerning the perceptual control of actions was introduced, based on the work of Arbib and Reichardt & Poggio amongst others. Within this framework the dynamical approach to motor control and coordination (Kelso, 1986; Kelso & Schönner, 1988; Kugler & Turvey, 1987; Saltzman & Kelso, 1987; Turvey & Kugler, 1984) is combined with computational and control theoretic approaches. It was argued that one should investigate the constraints imposed by the task an animal performs and determine, at an abstract “task-space” level (Saltzman & Kelso, 1987), the kinematics and dynamics of the motor system performing that task. This procedure means that one identifies the information that the system needs to perform the task and specifies the way that

information is used in the control of the action. A series of questions about perceptual control was specified: (1) what are the constraints the task itself places on the motor system and what are the basic informational requirements for performance of the task? (2) How is the task actually performed by a subject, in particular, what are the kinematic and dynamical descriptions of the performing system and what are the implications for the use of perceptual information by the system? (3) How is the information being used by the system available to the perceptual systems? (4) If the information is available from a variety of sources which of these are used by animal under study and how is this information extracted by the perceptual systems? Subsequent chapters investigated the control of timing in interceptive tasks and the perceptual information that could be used by human subjects to effect such control.

Chapter 3: here the task of intercepting a moving object was considered in some detail. It was argued that for tasks like catching and hitting, as performed by a human subject, effective timing of interception requires information about the time-to-contact of the moving object with the place where the interception is to be made. Accounts not involving the use of time-to-contact information due to Chapman (1968) and von Hofsten (1983) were examined and shown to be inadequate for effective timing in these tasks. The hypothesis due to Lee that effective timing does not require that the acceleration of the moving target be taken into account in computing time-to-contact information was considered.

It was concluded that time-to-contact information computed under the assumption that the relative acceleration of target and interceptor is zero may be sufficient for the timing of the interceptive actions a person is likely to perform.

Chapter 4: here the information available in the visual stimulus was analyzed. A review of the mathematics useful for analyzing information in a moving retinal image as developed by Longuet-Higgins and Prazdny (1980) was provided. This was used to show what time-to-contact information is visually available to a moving observer in a stationary environment under the assumption of zero relative acceleration, following the analysis of this problem provided by Lee (1980). The variable τ , introduced by Lee, was discussed and several different definitions of this vari-

able distinguished. In particular, a local tau variable (τ_L) was distinguished from a global one (τ_g). The problem of what information about the time-to-contact of a moving target with a moving observer and with a stationary observer is visually available was considered and two new mathematical results were derived showing how the time-to-contact of a moving target with any specifiable visual direction is visually available assuming zero relative acceleration.

Chapter 5: here the problems of extracting the time-to-contact information described in chapter 4 were analyzed. The problems associated with the use of τ_L were dealt with in detail and methods by which they might be overcome were suggested. The possibility of obtaining time-to-contact information with the use of two or more different perceptual systems was discussed and the notions of intermodal and multimodal timing information were defined.

Chapter 6: here existing empirical data was discussed. It was argued that it is not possible, in general, to account for the observed accuracy of human interceptive actions if one assumes that tau alone is the source of information about time-to-contact. It was concluded that more general time-to-contact information, perhaps that described in chapter 4, is typically used by human subjects. Much existing empirical work has sought to decide between tau or perceived distance of the target divided by its perceived velocity as the source of time-to-contact information by attempting to show that there is more to the perception of time-to-contact than tau alone can account for. It was argued that if it were found that there is more to time-to-contact perception than tau, this could not be taken as support for distance divided by velocity as the source of the information. The distance divided by velocity account needs to be fully justified and worked out if it is to constitute an empirically testable alternative to an account based on tau.

Chapter 7: the results of several experiments were presented which supported the hypothesis that both a moving target's rate of change of direction and position of the interception point are used in timing interceptions of targets which are not on a collision course with the observer's eyes. Such results are consistent with the use of the time-to-interception point information described in chapter 4. Although the results do not firmly establish that this information is actually the information used, they do contradict the alternative hypothesis that either local tau on its own or time-to-nearest approach information are sufficient to account for the timing

of interceptive actions. The results also support the idea, described in chapter 5, that timing information can be obtained using more than one perceptual system. The results show that although timing performance is both more accurate and more precise when visual information about the interception point and the rate of change of target direction is present, in the absence of visual information performance does not degrade to the level that would be expected were timing based only on the visual information remaining. This suggests that the rate of change of direction and interception point information can be obtained using articular proprioception as suggested in chapter 5.

In this final chapter, some of the empirical and theoretical implications of the work reported in this thesis are outlined and discussed.

§8.2. ADDITIONAL IMPLICATIONS AND FURTHER RESEARCH

The analysis presented in this thesis raises a number of empirical questions concerning the nature and use of time-to-contact information none of which have so far been explicitly addressed in the literature. In this section three questions will be briefly discussed and empirical methods for investigating them suggested.

What Timing Information is Used? As pointed out in chapter 6 no direct evidence that people use "direct" strategies rather than strategies based on prior computation of target distance and velocity to derive time-to-contact information to control the timing of interceptive actions has been provided. Various psychophysical experiments indicate that direct tau-based strategies can be used (Schiff & Detwiler, 1979; Simpson, 1988; Todd, 1981) and that image expansion is extracted by the human visual system (Regan, 1986; Regan & Beverley, 1978) but this evidence is not sufficient to allow the conclusion that direct tau-based strategies are used to obtain time-to-contact information when timing an interception. In chapter 6 it was argued that the distance÷velocity strategy provided an unconvincing alternative to direct strategies and needs to be more completely thought out, but no direct experimental evidence against its use in timing interceptions has been published.

Very recently, Savelsbergh, Whiting and Bootsma (1991) directly manipulated the rate of image expansion of a real approaching target. By controlled deflation or inflation of an approaching balloon image expansion could be made larger or

smaller than that which would be produced by simple approach of a rigid object. Using this method one would be able to show, as Savelsbergh et al. did, whether the movements involved in an interceptive action are dependent on the target's image expansion. However, it is not possible to conclude that a direct strategy rather than a distance÷velocity strategy is being used to obtain the information controlling the act's timing. Savelsbergh et al.'s experiment does establish, however, that relative rate of image expansion (or its reciprocal, local tau) is important in the control of the action. Nonetheless, as shown in chapter 6, relative rate of image expansion could be used to obtain information about a moving target's velocity and this velocity information could be used in conjunction with information about the target's distance to compute time-to-contact. Thus, use of relative rate of image expansion is not sufficient to distinguish direct from distance÷velocity strategies.

In chapter 6, we argued that the distance÷velocity strategy is a very unconvincing alternative to a direct strategy and that the former is perhaps impossible to distinguish empirically from the latter, at least in the case of direct collision approaches with the observation point. Nevertheless, this does indicate the difficulty involved in showing unequivocally that a certain source of information is used in the control of an action. Although we may be able to determine that certain perceptual variables are used in the control of an action, if these variables can be used in alternative ways to obtain the necessary control information, then simply showing that they are used is not sufficient to determine how the information is actually obtained. Thus, although the results of chapter 7 strongly suggest that a moving target's rate of change of direction and the position of the interception point are used to time interception of the target, they do not establish that the time-to-interception point information described in chapter 4 which involves these two variables is the information being used. Nor do they establish that the role of these two variables in interceptive timing is exactly the role they play in the time-to-interception point information. If this could be investigated, it would lead to a better understanding of the timing information actually used by subjects.

More quantitative experimental work is required to determine the precise role of the rate of change of target direction and the position of the interception point in interceptive timing. The best way to obtain more quantitative results would be to experimentally manipulate in a precisely controlled way the information specifying the target's rate of change of direction and the position of the interception point.

It is also important, as the results of the experiments in chapter 7 show, to employ a task that actually involves timing an interception. A task which involves actually grasping a moving target seems to give a smaller variable error than the trapping task used in experiments 2a and b as indicated by the results of Savelsbergh et al. who employed such a catching task. A grasping task is therefore preferable to the trapping task. One possible experiment might be as follows. Subjects would be required to grasp a self-luminous moving target at a fixed point in space as in Savelsbergh et al.'s experiment — there should be temporal but not spatial uncertainty in the task. Instead of fixed panels of LEDs providing peripheral visual information about the environment, one could use TV monitors which present computer controlled displays. The displays could then be animated during performance of the task so that subjects would be provided with non-veridical information about the rate of change of target direction — if the displays move in one way the visual information would be as if the eyes move faster than they actually do, if they move in the opposite way the visual information would be as if the eyes move slower than they actually do¹ (the right hand display moves in the opposite direction to the left hand display, of course). One therefore predicts a quantifiable effect on the timing error if the peripheral displays are providing visual information about the target's rate of change of direction. Timing should become later and later as a function of peripheral display velocity in one direction of motion and earlier and earlier as a function of display velocity in the opposite direction of motion.

Rotating Objects An effective experimental technique to determine how well the human visual system deals with non-spherically symmetric objects which rotate when they approach an observer would be to measure time-to-contact difference thresholds (Simpson, 1988; Todd, 1981) for such objects. The psychophysical procedure is fairly straightforward. Observers watch on a computer screen the simulated approach of two objects; a single frame of the display might look something like figure 8.1. The observer's task is, after watching a short animation of the object's approach, to indicate (perhaps by pressing an appropriate button) which of the two objects will reach him sooner (the task is forced-choice — one or other button must be pressed). The time-to-contact difference threshold is the differ-

¹ Note that Warren and colleagues (Warren & Hannon, 1988, 1990) have shown that in the absence of visual information about the fixed environment, displays which simulate eye rotation induce the perception of eye rotation even if the subjects eyes are not, in fact, moving.

ence between the two simulated objects' time-to-contacts at which the observer can correctly determine which will arrive sooner on some criterion percentage of trials (usually 75% or more). One arrives at this threshold using some standard psychophysical procedure. Although the implications of measured time-to-contact difference thresholds for the actual performance of interceptive actions is unclear, they are potentially able to supply useful information about the capabilities of the human visual system (see below) and it may be possible to use them to derive predictions about aspects of actual human performance which could then be tested. For example, Regan and Beverley (1979) derived predictions of the accuracy of intercepting moving targets based on their estimates of thresholds for image expansion and rate of change of disparity, though these were not checked against observer's performance of such interceptions.

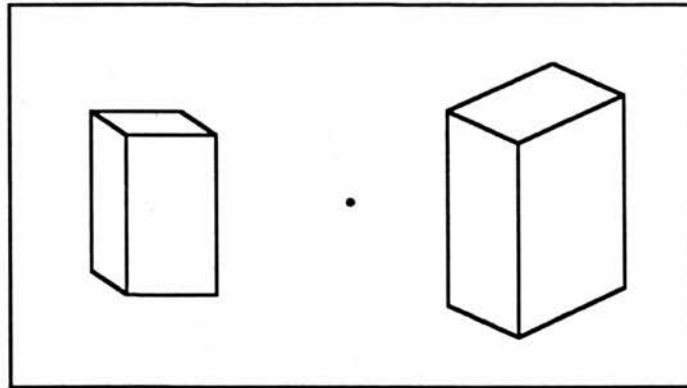


Figure 8.1. Single frame of a putative computer animation simulating the approach of two rectangular objects of identical shape and size (the smaller one is further away) directly towards the observer. The observer fixates the spot in the centre of the screen.

In the experiments of Todd (1981) and Simpson (1988) the changes in the size of the simulated approaching objects were only due to the object's approach. It is suggested that similar studies be run in which changes in size can be due to both approach and rotation. In these studies, both simulated objects (cf figure 8.1) would be the same shape and size and would be rotating at the same angular speed. Using this method it would be possible to explore the dependency of time-to-contact difference thresholds on object rotation. The following questions could be investigated:

- (1) Does the time-to-contact difference threshold change when there is object rotation about a single axis? If this case is more difficult to deal with than approach with no rotation thresholds should be higher.
- (2) Does the threshold depend on rate of rotation? If the visual system simply measures τ along an axis perpendicular to the axis of rotation, as suggested in chapter 5, then the threshold should be independent of rate of rotation.
- (3) How do time-to-contact difference thresholds change when there is rotation about two axes with both of these rotations causing changes in image size and shape? If a temporal averaging process is being used as suggested in chapter 5, it would be expected that viewing times in excess of the period of rotation are required for accurate performance.

Mechanisms of Time-to-contact Extraction Very little is known about the mechanisms which are responsible for the extraction of time-to-contact information. Koenderink and van Doorn (1976) pointed out that the optic flow (or, more precisely, the retinal projection of the optic flow) could be conveniently captured in the outputs of local mechanisms which responded to the expansion, rotation and deformation components of the flow (see figure 8.2) the outputs of such mechanisms should be independent of eye movements. Expansion detectors are of potential use in the computation of time-to-contact since τ is the reciprocal of the expansion. Regan and Beverley (1978) demonstrated psychophysically the existence of mechanisms specifically sensitive to changing retinal image size and called these mechanisms "looming detectors". What is not clear, however, is exactly what role looming or expansion detectors might play in the extraction of time-to-contact information. They cannot extract time-to-contact information directly as is sometimes suggested (e.g., Regan, 1986; Simpson, 1988) for at least the following three reasons:

- (1) The mechanisms proposed by Koenderink and Regan have a finite and constant retinal extent (receptive field size), whereas an approaching object's image changes size and cannot be expected to stimulate the same expansion detectors from one moment to the next (figure 8.3).
- (2) Required time-to-contact information will frequently require variables in addition to expansion information (τ) as discussed in chapters 4 and 6.

- (3) The mechanisms will respond to expansion regardless of its origins and so will respond even if the image is changing size due to object rotation or inflation/deflation.

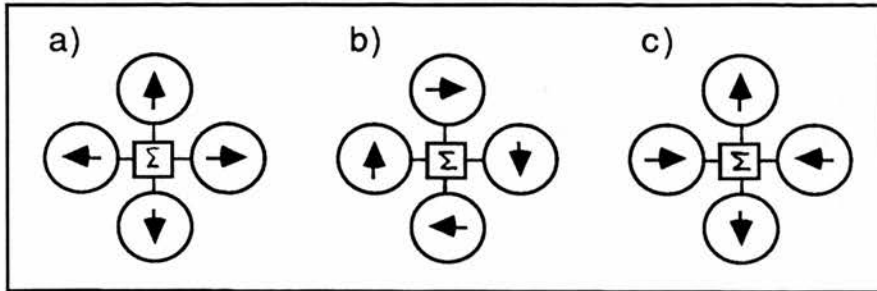


Figure 8.2 The construction of physiologically plausible a) divergence (expansion), b) curl (local rotation) and c) deformation (shear) detectors from directionally selective motion units (after Koenderink & van Doorn, 1976).

The implication is that the extraction of time-to-contact information is a complex process and cannot be done using expansion detectors alone. In Koenderink's scheme the three components of optical flow as extracted by the locally defined retinal mechanisms shown in figure 8.3 form the input to higher level mechanisms such as structure from motion processes (Koenderink, 1986; Koenderink & van Doorn, 1986). It is possible, therefore, that expansion detecting mechanisms could provide part of the input to processes which extract time-to-contact information even though they do not provide that information themselves.

This possibility was overlooked by Simpson (1988) who investigated how time-to-contact difference thresholds were affected by object rotations which led to rotations of the object's image but no change in its size or shape. He found that thresholds were raised when the image rotated and concluded that the rotatory component was not completely removed from the image motion (which was the composition of a rotatory component and a expansion component) in the judgement of time-to-contact. This led him to conclude that expansion detectors like those suggested by Koenderink were either not implicated in the extraction of time-to-contact information or if they were, then they were not true expansion detectors but responded to local rotation (or curl) as well. This would be a fair conclusion if it could be argued that expansion detectors were the sole source of time-to-contact information

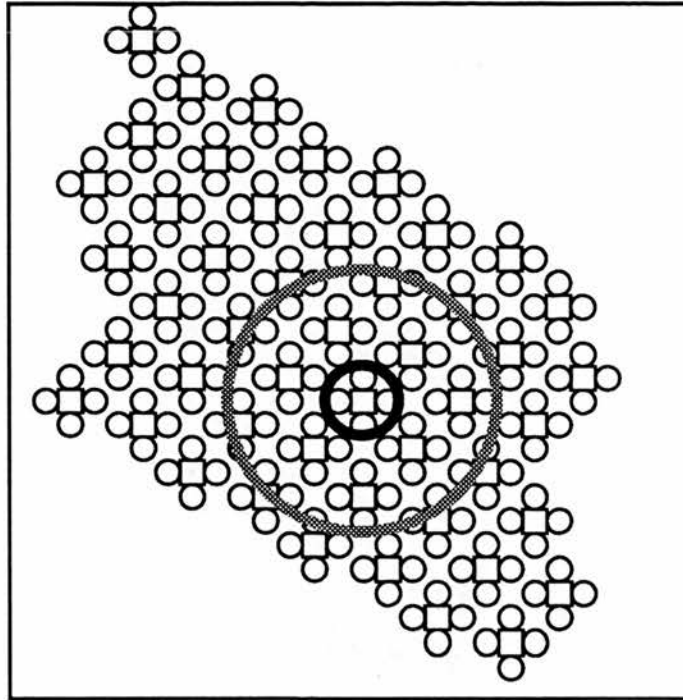


Figure 8.3. A schematic retinal surface is covered with an even pattern of expansion detectors. At one instant of time the image of the edge of an approaching object (black circle) stimulates a single expansion detector. At a later time the object's image (grey circle) is much larger and stimulates a number of different detectors.

in Simpson's experiments. However, as argued above, the outputs of expansion detectors alone cannot be used as the source of time-to-contact information but require higher levels of processing which interpret their output and combine it (if necessary) with other information derived from the perceptual input. Simpson's results could be due to these higher levels failing to deal effectively with target rotations in the image plane rather than to the absence or failures of expansion detectors. Thus, it cannot be concluded from these results whether or not expansion detectors of the type proposed by Koenderink play a role in time-to-contact perception.

To summarize, although expansion detectors may be implicated in the extraction of time-to-contact information they are clearly not the whole story. Their output is potentially useful if it can be related to the images of moving objects and surfaces. This may mean that some sort of segmentation process, which identifies different regions of the visual field as the images of different objects and surfaces, is

used in conjunction with the outputs of expansion detectors to compute the expansion of the images of moving objects. The effects of object rotation will need to be filtered out from such expansion information if useful time-to-contact information is to be obtained. Ways in which this might be achieved were discussed in chapter 5. It may be concluded that the mechanisms of time-to-contact perception in human vision are far from being understood. How well the visual system deals with the various cases described earlier needs to be thoroughly investigated, presumably psychophysically along the lines suggested, and models devised which are able to account for the data. Through the testing of such models against the performance of human subjects it may be hoped that the mechanisms of time-to-contact in the visual system will come to be better understood. As a simple example, the strategies for extracting time-to-contact information in the presence of object rotations described in chapter 5 could be investigated as accounts of human abilities using psychophysical methods as discussed at the end of the last section.

§8.3. CONCLUSIONS

It is clear from the analysis presented in this thesis that understanding of what time-to-contact information the human visual system uses in timing interceptive actions and how this information is extracted is very poorly understood. Many theoretical questions were examined, some of these have not been looked at before and those that have typically been given only a cursory treatment and have not been investigated empirically in a satisfactory way. Many empirical questions that emerged from the theoretical analysis were formulated precisely and experimental methods for investigating them were suggested. The results of some preliminary experiments were reported which lend some support to some of the theoretical results presented in chapters 4 and 5. It was found that although the results were consistent with use of the constant velocity strategy for interceptive timing, they were not consistent with use of tau alone nor with the use of time-to-nearest approach information. The results support the conclusion that the timing of interceptions of a moving target which bypasses the observer involves perception of the target's rate of change of direction and position of the interception point as would be predicted were subjects using the time-to-interception point information described in chapter 4.

The picture that is emerging from the work presented is that the perception

of time-to-contact is far more complicated than the simple picture that has found its way into the textbooks (e.g., Bruce & Green, 1985) in which tau defined as the reciprocal of the relative rate of a target's retinal image dilation is considered to be the perceptual source of time-to-contact information. Tau alone cannot, in general, account for the timing of interceptions which involve objects or target surfaces not on a collision path with the point of observation or objects whose rate of retinal image dilation is below threshold (e.g., in certain video games as considered in chapter 6) or non-spherically symmetric objects (e.g, rugby or American footballs) which rotate as they approach the observer. Different sources of time-to-contact information are appropriate under different circumstances: some sources involve the use of image expansion while others do not, and sometimes, when image expansion is required, a component due to target rotations needs to be removed from a component due to motion relative to the observer. In addition, some types of time-to-contact information can be obtained by human subjects using more than vision alone (other perceptual systems are implicated), and it will sometimes be necessary to obtain the information in this way (e.g., in the experimental tasks described in chapter 7).

It seems clear that there is far more to time-to-contact perception than tau. It remains to be seen whether human subjects switch between sources of time-to-contact information so as to obtain the most accurate estimates under different circumstances and whether they can filter out image expansion due to target rotation from image expansion due to target translation. Everyday experience would suggest that human observers do indeed do both of these things but this needs to be established experimentally — some possible experimental methods were outlined earlier in this chapter.

Appendix A

Accelerative Approaches

In this appendix expressions for time-to-contact information for some simple cases in which a target object is moving with constant linear acceleration are derived. Consider a spherical object on a collision course with a point of observation O . At an instant of time t let the distance between the object and O be $Z(t)$, the relative speed be $V(t)$ and signed magnitude of the relative acceleration $A(t)$ (assumed constant). Let the rate of change of direction of the object relative to O at time t be zero. Dependency on t will be suppressed in what follows. Assuming that the expression $Z/V = \tau_L$ is valid it may be differentiated with respect to time to give

$$\dot{\tau}_L = \frac{-V^2 - ZA}{V^2}, \quad (A1)$$

Lee, Young, Reddish, Lough & Clayton, (1983) used the following standard formula for distance travelled under constant acceleration to obtain an expression for T_c during linear accelerative approach:

$$Z = VT_c + \frac{1}{2}AT_c^2. \quad (A2)$$

Treating this as a quadratic in T_c and taking the positive root, the expression derived by Lee et al. for the time-to-contact of the moving object with O under the conditions described above is

$$T_c = \frac{\tau_L[1 - \sqrt{(-2\dot{\tau}_L - 1)}]}{(1 + \dot{\tau}_L)}. \quad (A3)$$

This shows that, for constant accelerative approach, time-to-contact is optically specified by an expression considerably more complicated than τ_L alone which specifies T_c in the same circumstances but with V constant.

Now consider the geometry illustrated in figure A1. Assuming constant object velocity, following the arguments presented in chapter 4 one can apply the sine rule to obtain (suppressing dependency on t as usual)

$$\frac{R}{\sin \beta} = \frac{S}{\sin \psi} \quad (A4)$$

Differentiating A4 with respect to time and substituting T_p (the time-to-contact of the moving object with the point p) for $-S/\dot{S}$ and τ_L for $-R/\dot{R}$ and denoting

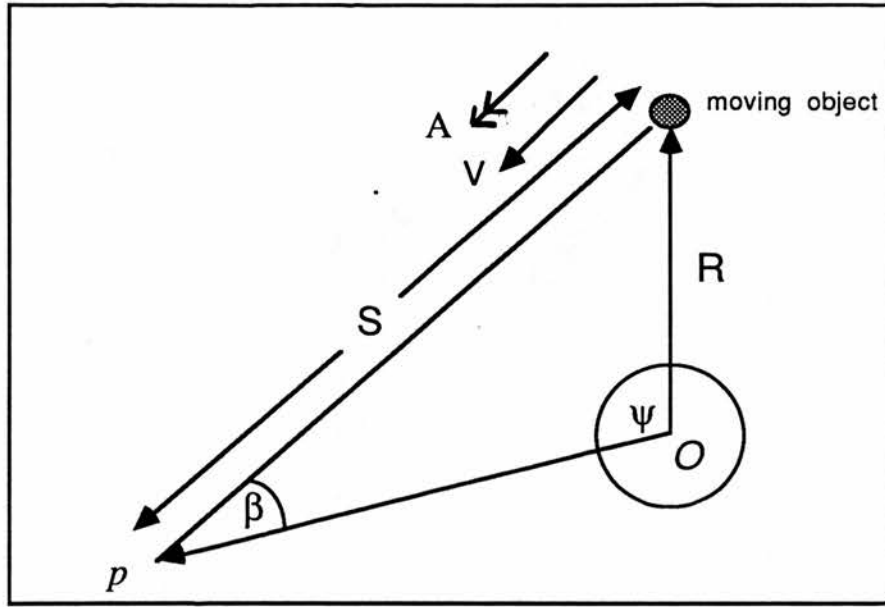


Figure A1. Geometry of time-to-contact. A spherical object (e.g., a ball) is moving in the plane of the page towards the point p past the centre of projection O (the angle β is constant). The object's instantaneous velocity is V (considered constant in the derivation of equation A5) and it has constant linear acceleration, A . The circle represents a slice through a spherical imaging surface centred on O .

$-\dot{\psi}$ by u (the object's rate of change of direction as seen from O) yields, with rearrangement, the expression

$$T_p = \frac{\tau_L}{1 + \tau_L u \cot \psi}. \tag{A5}$$

This expression is equation 4.23 in chapter 4 and shows that the T_c with any point p on the upcoming path of the moving object is specified in terms of variables potentially available from the optic projection. Relaxing the constant velocity constraint means that equation A4 can be differentiated twice and simplified to yield

$$\frac{2V}{S} - 2\dot{\psi} \cot \psi - \frac{\dot{\psi}^2 S}{V} - \frac{2\dot{R}}{R} + \frac{2S\dot{R}\dot{\psi} \cot \psi}{VR} + \frac{S\ddot{\psi} \cot \psi}{V} + \frac{S\dot{R}\ddot{R}}{V\dot{R}\ddot{R}} = \frac{A}{V}. \tag{A6}$$

Where V is the instantaneous speed of the object. Notice that the quantity $-\dot{R}/R$ is the reciprocal of τ_L (which will be denoted ρ) and that S/V is T_p for which an expression in terms of optic variables has already been derived (equation A5). An expression for \ddot{R}/\dot{R} in terms of τ_L and $\dot{\tau}_L$ can be obtained by differentiating $\tau_L = -R/\dot{R}$ with respect to time and substituting for $-R/\dot{R}$ in the resulting expression. This yields the expression, $-\ddot{R}/\dot{R} = (\dot{\tau}_L + 1)/\tau_L$. Equation A6 can be

rewritten completely in terms of optic variables as follows (writing T_p for the right hand side of A5):

$$\frac{A}{V} = \frac{2}{T_p} - 2\dot{\psi} \cot \psi - \dot{\psi}^2 T_p + 2\varrho - 2\varrho T_p \dot{\psi} \cot \psi + T_p \ddot{\psi} \cot \psi - T_p \varrho^2 (\dot{\tau}_L + 1). \quad (A7)$$

To obtain an expression for time-to-contact under accelerative approach one finds an expression for T_c from equation A2 by applying the standard formula for the roots of a quadratic, noting that in this case T_c is interpreted as the time-to-contact under accelerative approach with the point p . Thus,

$$T_c = \frac{-1 + \sqrt{1 + 2(A/V)(S/V)}}{(A/V)}. \quad (A8)$$

Since expressions for both A/V and S/V have been derived giving these quantities in terms of optic variables (equations A7 and A5 respectively) the time-to-contact of the moving object with any point p is specified optically under conditions of constant accelerative approach. It is clear that the relevant expression is a complex relation (it is not written out here but if desired can be obtained by simple substitution into A8) between several optic quantities (two of these, $\ddot{\psi}$ and $\dot{\tau}_L$, being second temporal derivatives) and is considerably more complicated than the corresponding expression which assumes constant velocity (equation A5).

REFERENCES

- Abraham, R. & C. D. Shaw (1982) Dynamics: The Geometry of Behaviour. Volume 2. Santa Cruz: Ariel Press.
- Alderson, G. K. W., Sully, H. & D. Sully (1974) An operational analysis of a one-handed catching task using high speed photography. Journal of Motor Behavior, **6**, 217-226.
- Aloimonos (1988) Unifying texture and shading through an active observer. Proceedings of the Royal Society of London, **B238**, 25-37.
- Amit, D. J. (1989) Modelling Brain Function. The World of Attractor Neural Networks. Cambridge: Cambridge University Press.
- Anand, D. K. (1984) Introduction to Control Systems (2nd edition). Oxford: Pergamon Press.
- Anderson, G. J. (1986) Perception of self-motion: psychophysical and computational approaches. Psychological Bulletin, **99**, 52-65.
- Arbib, M. A. (1972) The Metaphorical Brain. Wiley Interscience.
- Arbib, M. A. (1981) Perceptual structures and distributed motor control. In V. Brooks ed., Handbook of Physiology: The nervous System, Vol II Motor Control, Bethesda, American Physiological Association.
- Arbib, M. A. (1987) Levels of modelling of mechanisms of visually guided behaviour, Behavioural & Brain Sciences, **10**, 407-465.
- Arbib, M. A. (1988) Brains, Machines, and Mathematics (2nd edition). New York: Springer-Verlag.
- Arbib, M. A. & R. Hanson (1987) Introduction. In M. Arbib & R. Hanson (eds.), Vision, Brain & Cooperative Computation. Cambridge Mass.: M.I.T. Press.
- Ashby, W. R. (1956) Design For A Brain. London: Chapman & Hall.
- Ballard, D. H. (1986) Cortical connections and parallel processing. Behavioural & Brain Sciences, **9**, 67-120.
- Ballard, D. H., Sejnowski, T. & G. E. Hinton (1983) Parallel visual computation. Nature, **306**, 21-26.
- Barnett, S. & R. G. Cameron (1985) Introduction to Mathematical Control Theory (2nd edition). Oxford: Oxford University Press.
- Barwise, J. & J. Perry (1983) Situations and Attitudes. Bradford Books: MIT Press.

- Berkinblit, M. B., Feldman, A. G. & O. Fukson (1986) Adaptability of innate motor patterns and motor control mechanisms. Behavioural and Brain Sciences, 9, 585-638.
- Bernstein, N. A. (1967) The Coordination and Regulation of Movements. Oxford: Pergamon Press.
- Bizzi, E. (1980) Central and Peripheral mechanisms in motor control. In G. Stelmach & J. Requin (ed.) Tutorials in Motor Behaviour. Amsterdam: North Holland.
- Bizzi, E., Accornero, N., Chapple, W., & N. Hogan, (1982) Arm trajectory formation in monkeys. Experimental Brain Research, 46, 139-143.
- Bizzi, E., Accornero, N., Chapple, W., & N. Hogan, (1984) Posture control and trajectory formation during arm movement. Journal of Neuroscience, 4, 2738-2744.
- Bootsma, R. J. (1989) Accuracy of perceptual processes subserving different perception-action systems. Quarterly Journal of Experimental Psychology, 41A, 489-500.
- Bootsma, R. J. & van Weiringen (1989) Timing an attacking forehand drive in table tennis. Journal of Experimental Psychology: Human Perception & Performance, 16, 21-29
- Brady M., J. Hollerbach, T. Johnson, T. Lozano-Perez & M. Mason (1982) Robot Motion: Planning and Control. Cambridge Mass: MIT Press
- Braemer, W. & H. Braemer (1958) Orientation of fish to gravity. Limnology & Oceanography, 3, 362-372.
- Bruce, V. & P. Green (1985) Visual Perception: Psychology, Physiology and Ecology. Hillsdale N.J. : Lawrence Erlbaum.
- Carello, C., A. Groszofsky, F. Reichel, F., H.Y. Soloman, & M. T. Turvey (1989). Reaching. Ecological Psychology, 1, 27-54.
- Cavallo, V. & M. Laurent (1988) Visual information and skill level in time-to-collision estimation. Perception, 17, 623 -632.
- Cavallo, V., O. Laya & M. Laurent (1986) The estimation of time-to-collision as a function of visual stimulation. In A. G. Gale, M. Freeman, C. Haselgrave, P. Smith & S. Taylor (ed.) Vision In Vehicles. Amsterdam: Elsevier.
- Chater, N. J. (1990) Information and information Processing. Unpublished PhD Thesis, University of Edinburgh.
- Chapman (1968) Catching a baseball. American Journal of Physics, 36, 868-870.
- Clocksini, W. F. (1980) Perception of surface slant and edge labels from optical flow: a computational approach. Perception, 9, 253-269.

- Cutting, J. E. (1986) Perception with an Eye for Motion. London: Bradford Books, M.I.T. Press.
- Diggles, V. A., M. Grabiner & J. Garhammer (1988) Skill level and efficacy of effector visual feedback in ball catching. Perceptual and Motor Skills, **64**, 987-993.
- Doner, J., J. Lappin & G. Peretto (1984) Detection of three-dimensional structure in moving optical patterns. Journal of Experimental Psychology: Human Perception and Performance, **10**, 1-11.
- Dretske, F. I. (1981) Knowledge and the Flow of Information. London: Bradford Books, M.I.T. Press.
- Feldman, A. G. (1980a) Superposition of motor programs. I. Rhythmic forearm flexion in man. Neuroscience, **5**, 91-95.
- Feldman, A. G. (1980b) Superposition of motor programs. II. Rapid forearm flexion in man. Neuroscience, **5**, 91-95.
- Feldman, A. G. (1986) Once more on the equilibrium-point hypothesis (λ model) of motor control. Journal of Motor Behaviour, **18**, 17-54.
- Feldman, A. G. & M. L. Latash (1982) Interaction of afferent and efferent signals underlying joint position sense: empirical and theoretical approaches. Journal of Motor Behaviour, **14**, 174-193.
- Feynman, R., R. Leyton & M. Sands (1964) The Feynman Lectures on Physics, Vol II. London: Addison-Wesley.
- Fischer, M. H. & A. Kornmuller (1930) Optokinetisch ausgeloste bewegungswahrnehmung und optokinetischer nystagmus. Journal Psychologische Neurologie (Leipzig), **41**, 273-308.
- Fischman, M. G. & T. Schneider (1985) Skill level, vision and proprioception in simple one-handed catching. Journal of Motor Behaviour, **17**, 219-229.
- Fodor, J. A. (1980) Methodological solipsism considered as a research strategy cognitive psychology. Brain and Behavioural Sciences, **3**, 63-110.
- Fodor, J. A. & Z. W. Pylyshyn (1981) How direct is visual perception: some comments on Gibson's ecological approach. Cognition, **28**, 3-71
- Fraser, J. (1908) A new illusion of visual direction. British Journal of Psychology, **2**, 307-320.
- Gibson, E. J. (1969) Principles of Perceptual learning and Development. New York: Prentice Hall.
- Gibson, J. J. (1950) The Perception of the Visual World. Boston: Houghton-Mifflin.
- Gibson, J. J. (1958) Visually controlled locomotion and visual orientation in animals. British Journal of Psychology, **49**, 182-194.
- Gibson, J. J. (1966) The Senses Considered as Perceptual Systems. Boston: Houghton-Mifflin.

- Gibson, J. J. (1977) The theory of affordances. In R. E. Shaw & J. Bransford (eds.), Perceiving, Acting, Knowing: Towards an Ecological Psychology. Hillsdale N. J.: Lawrence Erlbaum.
- Gibson, J. J. (1979) The Ecological Approach to Visual Perception. Boston: Houghton-Mifflin.
- Gibson, J. J., P. Olum & F. Rosenblatt (1955) Parallax and perspective during aircraft landings. American Journal of Psychology, **68**: 372-385.
- Goodale, M. A., D. Pelisson & C. Prablanc (1986) Large displacements in visually guided reaching do not depend on vision of the hand or perception of target displacement. Nature, **320**, 748-750.
- Gottsdanker, R. M. (1956) The ability of human operators to detect the acceleration of target motion. Psychological Bulletin, **53**, 477-487.
- Gottsdanker, R. M., J. Frick & R. Lockerd (1961) Identifying the acceleration of visual targets. British Journal of Psychology, **52**, 31-42.
- Gregory, R. L. (1966) Eye and Brain. London: Weidenfeld & Nicholson.
- Gregory, R. L. (1972) Seeing as thinking. Times Literary Supplement, (June 23), 707-708.
- Gregory, R. L. (1974) The Mechanisms of Perception. London: McGraw-Hill.
- Groeger, J. A. & I. Brown (1988) Motion perception is not direct with indirect viewing systems. In A. G. Gale, M. Freeman, C. Haselgrave, P. Smith & S. Taylor (eds.) Vision In Vehicles. Amsterdam: Elsevier.
- Grossberg, S. (1976) Adaptive pattern classification and universal recoding II. Biological Cybernetics, **23**, 187-202.
- Grossberg, S. (1980) Direct perception and adaptive resonance (commentary on Ullman, 1980). Behavioural & Brain Sciences, **3**, 385.
- Grossberg, S. (1983) The quantized geometry of visual space. Behavioural & Brain Sciences, **6**, 625-692.
- Haken, H. (1983) Synergetics (3rd edition). Berlin: Springer Verlag
- Harel, D. (1987) Algorithmics: the Spirit of Computing. Addison Wesley.
- Hildreth, E. C. (1984) The Measurement of Visual Motion. Cambridge Mass.: M.I.T press.
- Hildreth, E. C. & C. Koch (1987) The analysis of visual motion: from computational theory to neuronal mechanisms. Annual Review of Neuroscience, **10**, 37-72.
- Hinton, G. E. (1980) Inferring the meaning of direct perception, Behavioural & Brain Sciences, **3**, 387-388.

- Hofsten, C. von (1983) Catching skills in infancy. Journal of Experimental Psychology: Human Perception and Performance, **9**, 75-85.
- Hofsten, C. von (1987) Catching. In H. Heuer & A. F. Sanders (ed.) Perspectives on Perception and Action. Hillsdale N. J.: Lawrence Erlbaum.
- Hofsten, C. von & D. N. Lee (1985) Dialogue on perception and action. In W. H. Warren & R. E. Shaw (ed), Persistence and Change. Hillsdale N. J.:Lawrence Earlbaum.
- Hogan, N. (1985) Impedance control: an approach to manipulation: part 1--theory. Journal of Dynamic Systems, Measurement and Control, **107**, 1-7.
- Hollerbach, J. M. (1982) Computers, brains and the control of movement. Trends in Neurosciences, June 1982, 189-192.
- Horn, B. K. P. (1974) Determining lightness from an image. Computer Graphics and Image Processing, **3**, 111-199.
- Horn, B. K. P. (1986) Robot Vision. Cambridge & New York: M.I.T. press.
- Howard, I. P. (1982) Human Visual Orientation. Chichester: John Wiley & Sons.
- Ingle, D. & Shook (1985) Action oriented approaches to visuo-spatial brain functions. In In D. Ingle, M. Jeannerod & D. Lee (eds), Brain Mechanisms and Spatial Vision. Dordrecht: Martinus Nijhoff.
- Jeannerod, M. (1988) The Neural and Behavioural Organization of Goal-Directed Movements. Oxford: Clarendon Press.
- Johansson, G. (1975) Visual motion perception. Scientific American: 76-88.
- Johansson, G. (1975) Studies on the visual perception of locomotion. Perception, **6**, 365-376.
- Jones, R. K. & D. N. Lee (1981) Why two eyes are better than one: the two views of binocular vision. Journal of Experimental Psychology: Human Perception and Performance, **7**, 30-40.
- Judge, S. J. & C. M. Bradford (1988) Adaptation to telestereoscopic viewing measured by one-handed catching performance. Perception, **17**, 783-802.
- Kalman, R. E., Falb, P. & M. A. Arbib (1969) Topics in Mathematical System Theory. London: McGraw-Hill.
- Kelso, J. A. S. (1981) The rhythmical basis of movement. Bulletin of the Psychonomic Society, **18**, 63-69.
- Kelso, J. A. S. (1986) Pattern formation in speech and limb movements involving many degrees of freedom. In Heuer, H. & C. Fromm, (ed.), Generation and Modulation of Action Patterns. New York: Springer-Verlag

- Kelso, J. A. S. & G. Schoner (1988) Towards a physical (synergetic) theory of biological coordination. In R. Graham (ed.), Lasers and Synergetics (Springer series in physics vol 8). Berlin: Springer-Verlag.
- Koenderink, J. J. (1980) Why argue about direct perception? (Commentary on Ullman, 1980). Behavioural & Brain Sciences, **3**, 390-391.
- Koenderink, J. J. (1985) Space, form and optical deformations. In D. Ingle, M. Jeannerod & D. Lee (eds), Brain Mechanisms and Spatial Vision. Dordrecht: Martinus Nijhoff.
- Koenderink, J. J. (1986) Optic flow. Vision Research, **26**, 161-180.
- Koenderink, J. J. & A. J. van Doorn (1975) Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. Optica Acta, **22**, 773 -791.
- Koenderink, J. J. & A. J. van Doorn (1976) Local structure of the movement parallax of the plane. Journal of the Optical Society of America, **66**, 717-723.
- Koenderink, J. J. & A. J. van Doorn (1977) How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow. In G. Hauske & E. Butenandt (ed.) Kybernetik 1978. Oldenberg: Muenchen.
- Koenderink, J. J. & A. J. van Doorn (1981) Exterosppecific component of the motion parallax field. Journal of the Optical Society of America, **71**, 953-957.
- Koenderink, J. J. & A. J. van Doorn (1987) Facts on optic flow. Biological Cybernetics, **56**, 255-260.
- Kugler, P. N. & M. T. Turvey (1987) Information, Natural Law and the Self-Assembly of Rhythmic Movement. Hillsdale N. J.: Lawrence Erlbaum.
- Land, E. H. (1986) Recent advances in retinex theory. Vision Research, **26**, 7-21.
- Land, E. H. & J. J. MacCann (1971) Lightness and retinex theory. Journal of the Optical Society of America, **61**, 1-11.
- Lawton, D., J. Rieger & M. Steenstrup (1987) Computational aspects of visual motion processing. In M. Arbib & R. Hanson (ed.), Vision, Brain and Cooperative Computation. Cambridge Mass.: M.I.T. Press.
- Laurent, M., R. Dinh Phung, & H. Ripoll (1989). What visual information is used by riders in jumping? Human Movement Science, **8**, 481-501.
- Lee, D. N. (1974) Visual information during locomotion. In R. Macleod & H. Pick (ed.), Perception: Essays in honour of J. J. Gibson. Ithaca, N.Y.: Cornell University Press.

- Lee, D. N. (1976) A theory of visual control of braking based on information about time-to-collision. Perception, **5**, 437-459.
- Lee, D. N. (1978) The functions of vision. In H. Pick & E. Saltzman (ed.) Modes of Perceiving and Processing Information. Hillsdale NJ: Lawrence Erlbaum.
- Lee, D. N. (1980a) The optic flow field: the foundation of vision. Philosophical Transactions of the Royal Society of London, **B290**, 169-179.
- Lee, D. N. (1980b) Visuo-motor coordination in space-time. In G. E. Stelmach & J. Requin (eds), Tutorials in Motor Behaviour. Amsterdam: North Holland.
- Lee, D. N. (1990) Getting around with light or sound. In R. Warren and A. Wertheim (ed.), The Perception and Control of Self-Motion. Hillsdale, NJ: Lawrence Erlbaum.
- Lee, D. N. & E. Aronson (1974) Visual proprioceptive control of standing in human infants. Perception & Psychophysics, **15**: 529-532.
- Lee, D. N. & J. R. Lishman (1977) Visual control of locomotion. Scandinavian Journal of Psychology, **18**, 224-230.
- Lee, D. N., Lishman, J. R. & J. Thompson (1982) Regulation of gait in long jumping. Journal of Experimental Psychology: Human Perception and Performance, **8**, 448-459.
- Lee, D. N. & P. E. Reddish (1981) Plummeting gannets: a paradigm of ecological optics. Nature, **293**, 293-294.
- Lee, D. N. & D. S. Young (1985) Visual timing of interceptive action. In D. Ingle, M. Jeannerod & D. N. Lee (eds), Brain mechanisms and spatial vision. Dordrecht: Martinus Nijhoff.
- Lee, D. N., D. S. Young, P. E. Reddish, S. Lough & T. Clayton (1983) Visual timing in hitting an accelerating ball. Quarterly Journal of Experimental Psychology, **35A**, 333-346.
- Lishman, J. R. & D. N. Lee (1973) The autonomy of visual kinaesthesia. Perception, **2**, 287-294.
- Longuet-Higgins, H. C. (1986) Visual motion ambiguity. Vision Research, **26**, 181-183.
- Longuet-Higgins, H. C. & K. Prazdny (1980) The interpretation of a moving retinal image. Proceedings of the Royal Society of London, **B208**, 385 -397.
- Luenberger, (1979) Introduction to Linear Systems Theory. Addison Wesley.
- McLeod, P., C. McLaughlin & P. Nimmo-Smith (1985) Information encapsulation and automaticity. Evidence from the visual control of finely timed actions. In M. Posner & O. Marin (ed.) Attention & Performance XI. Hillsdale N. J.: Lawrence Erlbaum.

- McLeod, R. W. & H. E. Ross (1983) Optic flow and cognitive factors in time-to-collision estimates. Perception, **12**, 417-423.
- Mach, E. (1875) The Analysis of Sensations. English translation 1962, Dover, New York.
- Marr, D. C. (1982) Vision. San Francisco: W. H. Wiley.
- Marr, D. & T. Poggio (1976) Cooperative computation of stereo disparity. Science, **194**, 283-287.
- Marr, D. & T. Poggio (1979) A computational theory of human stereo vision. Proceedings of the Royal Society of London, **B204**, 151-180.
- Mayhew, J. E. W. (1983) Stereopsis. In O. Braddick & A. Sleight (ed.) Physical and Biological Processing of Images. Springer-Verlag.
- Mayhew, J. E. W. & J. Frisby (1982) Psychophysical and computational studies towards a theory of human stereopsis. Artificial Intelligence, **17**, 349-385.
- Meijer, O. & K. Roth (1988) (Eds.) Complex Movement Behaviour: the "Motor-Action" Controversy. Amsterdam: North Holland.
- Micheals, C. F. & C. Carello (1981) Direct Perception. New York: Prentice Hall.
- Morgan, M. J. & W. Hotopf (1989) Perceived diagonals in grids and lattices. Vision Research, **29**, 1005-1015.
- Morgan, M. J. & B. Moulden (1986) The Munsterberg figure and twisted cords. Vision Research, **26**, 1793-1800.
- Nakayama, K. (1985) Biological image motion processing: a review. Vision Research, **25**, 625-660.
- Nakayama, K. & J. M. Loomis (1974) Optical velocity patterns, velocity sensitive neurons and space perception: a hypothesis. Perception, **3**, 63-80.
- Oatley, K. (1978) Computational metaphor for perception. In Perception: Course in Cognitive Psychology (Part 2). Milton Keynes: Open University Press.
- Olson, H. F. (1958) Dynamical Analogies. San Francisco: Freeman.
- Pelisson, D., C. Prablanc, M. Goodale & M. Jeannerod (1986) Visual control of reaching movements without vision of the limb II. Experimental Brain Research, **62**, 303-311.
- Poggio, T. & C. Koch (1985) Ill-posed problems in early vision: from computational theory to analogue networks. Proceedings of the Royal Society of London, **B226**, 303-323.
- Poggio, T. & W. Reichardt (1976) Visual control of orientation behaviour in the fly. Part II. Quarterly Reviews of Biophysics, **9**, 376-438.

- Prazdny, K. (1980) How wrong was Gibson? (commentary on Ullman, 1980). Behavioural & Brain Sciences, **3**, 394-395.
- Prazdny, K. (1980) Egomotion and relative depth map from optical flow. Biological Cybernetics, **36**, 87-102.
- Prazdny, K. (1981) Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer. Computer Graphics and Image Processing, **17**, 238-248.
- Reed, E. S. (1982) An outline of a theory of action systems. Journal of Motor Behaviour, **14**, 98-134.
- Reed, E. S. (1984) From action gestalts to direct action. In H. T. A. Whiting (ed.), Human Motor Actions: Bernstein Reassessed. Amsterdam: North Holland.
- Reed, E. S. (1988) Applying the theory of action systems. In O. Meijer & K. Roth (eds.), Complex Movement Behaviour: The Motor-Action Controversy. Amsterdam: North Holland.
- Regan, D. M. (1986) Visual processing of four types of relative motion. Vision Research, **26**, 127-145.
- Regan, D. M. & K. Beverley (1978) Looming detectors in the human visual pathway. Vision Research, **18**, 415-421.
- Regan, D. M. & K. Beverley (1979) Binocular and Monocular stimuli for motion in depth. Changing-disparity and changing-size feed the same motion-in-depth stage. Vision Research, **19**, 1331-1342.
- Regan, D. M. & K. Beverley (1982) How do we avoid confounding the direction we are looking with the direction we are moving? Science, **215**, 194-196.
- Reichardt, W. (1986) Processing of optical information by the visual system of the fly. Vision Research, **26**, 113-125.
- Reichardt, W. & T. Poggio (1976) Visual control of orientation behaviour of the fly. Part I. Quarterly Reviews of Biophysics, **9**, 311-375.
- Reichardt, W. & T. Poggio (1981) Visual control of flight in flies. In W. Reichardt & T. Poggio (eds.), Theoretical Approaches in Neurobiology. Cambridge Mass.: MIT Press.
- Reichardt, W. & H. Wenking (1969) Optical detection and fixation of objects by fixed flying flies (*musca domestica*). Naturwissenschaften, **56**, 424-425.
- Rieger, J. H. (1983) Information in optical flows induced by curved paths of observation. Journal of the Optical Society of America, **73**, 339-344.
- Rosen, R. (1985) Anticipatory Systems. Oxford: Pergamon Press.
- Rosengren, K., H. Pick & C. von Hofsten (1988) The role of visual information in ball catching. Journal of Motor Behaviour, **20**, 150-164.

- Runeson, S. (1974) Constant velocity-not perceived as such. Psychological Research, **37**, 3-23.
- Runeson, S. (1975) Visual prediction of collision with natural and unnatural motion functions. Perception & Psychophysics, **18**, 261-266.
- Runeson, S. (1977) On the possibility of smart perceptual mechanisms. Scandinavian Journal of Psychology,
- Saltzman, E. (1979) Levels of sensorimotor representation. Journal of Mathematical Psychology, **20**, 91-163.
- Saltzman, E. & J. A. S. Kelso (1987) Skilled actions: a task dynamic approach. Psychological Review, **94**, 84-106.
- Savelsbergh, G. J., H. T. A. Whiting & R. J. Bootsma (1989) An ecological perspective on one-handed catching. In Proceedings of the International Congress of Psychology, Sydney, Australia.
- Savelsbergh, G. J., H. T. A. Whiting & R. J. Bootsma (1991) Grasping tau. Journal of Experimental Psychology: Human Perception and Performance, **17**, 315-322.
- Saxberg, B. V. H. (1987) Projected free fall trajectories I: theory and simulation. Biological Cybernetics, **56**, 159-175.
- Scheffe, H. (1953) A method for judging all contrasts in the analysis of variance. Biometrika, **40**, 87-104.
- Schiff, W. & M. L. Detwiler (1979) Information used in judging impending collision. Perception, **8**, 647-658.
- Schiff, W. & R. Oldak (1990) Accuracy of judging time to arrival: effects of modality, trajectory and gender. Journal of Experimental Psychology: Human Perception and Performance, **16**, 303-316.
- Schone, H. (1984) Spatial Orientation. Princeton, N.J. : Princeton University Press.
- Searle, J.R. (1983) Intentionality. Cambridge University Press.
- Sedgwick, (1973) The visible horizon: a potential source of visual information for the perception of size and distance. Dissertation Abstracts International, **34**, 1301B-1302B. (University microfilms no. 73-22530).
- Shannon, C. E. & W. Weaver (1949) Mathematical Theory of Communication. Urbana: University of Illinois Press.
- Sharp, R. H. & H. T. A. Whiting (1974) Exposure and occluded duration effects in a ball-catching skill. Journal of Motor Behaviour, **6**, 139-147.

- Sharp, R. H. & H. T. A. Whiting (1975) Information processing and eye movement behaviour in a ball-catching skill. Journal of Human Movement Studies, **1**, 124-131.
- Shaw, R. E. & J. Bransford (1977) Introduction. In R. Shaw & J. Bransford (ed.) Perceiving, Acting, Knowing: Towards an Ecological Psychology. Hillsdale NJ: Lawrence Erlbaum.
- Shaw, R. E. & M. T. Turvey (1981) Coalitions as models for ecosystems. In J. Kubovy & J. R. Pomerantz (ed.), Perceptual Organisation. Hillsdale N. J.: Lawrence Erlbaum.
- Simpson, W. (1988) Depth discrimination from optic flow. Perception, **17**, 497-512.
- Smyth, M. & A. M. Marriott (1982) Vision and proprioception in simple catching. Journal of Motor Behaviour, **14**, 143-152.
- Soloman, H. Y. (1988) Movement produced invariants in haptic exploration. Human Movement Science, **7**, 201-223
- Soloman, H. Y., C. Carello & M. T. Turvey (1984) Flow fields: the optical support for skilled action. In W. Strauss & J. Williams (ed.), Cognitive Sport Psychology. New York: Lansing.
- Stange, G. (1972) Die zentralnervöse Verrechnung optischer Afferenzen bei der Gleichgewichtshaltung von fischen. Journal of Comparative Physiology, **80**, 95-118.
- Stoffregen, T. A. & G. Riccio (1988) An ecological theory of orientation and the vestibular system. Psychological Review, **95**, 3-14.
- Subbarao, M. (1988) Interpretation of Visual Motion: A Computational Study. London: Pitman.
- Tenenbaum, J. M. & H. Barrow (1976) Experiments in interpretation guided segmentation. Stanford Research Institute Tech. Note 123.
- Thompson, W. T. (1983) Theory of Vibration with Applications (2nd Edition). London: George Allen & Unwin.
- Todd, J. T. (1981) Visual information about moving objects. Journal of Experimental Psychology: Human Perception and Performance, **7**, 795-810.
- Todd, J. T. (1985) Perception of structure from motion: is projective correspondence of moving elements a necessary condition? Journal of Experimental Psychology: Human Perception & Performance, **11**, 689-710.
- Torrey, C. (1985) Visual flow and the direction of locomotion. Science, **277**, 1064.
- Turvey, M. T. (1977) Preliminaries to a theory of action with reference to vision. In R. E. Shaw & J. D. Bransford (eds.), Perceiving, Acting, Knowing: Towards and Ecological Psychology. Hillsdale, N.J.: Lawrence Erlbaum.

- Turvey, M. T. (1986) Commentary on K.M. Sayre, Behavioural & Brain Sciences, **9**, 153-155.
- Turvey, M. T. & C. Carello (1986) The ecological approach to perceiving-acting. Acta Psychologica, **63**, 133-155.
- Turvey, M. T. & P. N. Kugler (1984) The ecological approach to perceiving and acting. In H. T. A. Whiting (ed) Human Motor Actions: Bernstein Reassessed. Amsterdam: North Holland.
- Turvey, M. T., Shaw, R. E. & W. Mace (1978) Issues in the theory of action. In J. Long & A. Badaeley (eds.), Attention and Performance IX. Hillsdale N. J.: Lawrence Erlbaum.
- Turvey, M. T., Shaw, R. E., Reed, E. S. & W. Mace (1981) Ecological laws of perceiving and acting: in reply to Fodor & Pylyshyn. Cognition, **9**, 237-304.
- Tyldesley, D. A. & H. T. A. Whiting (1975) Operational timing. Journal of Human Movement Studies, **1**, 172-177.
- Ullman, S. (1979) The Interpretation of Visual Motion. Cambridge Mass.: M.I.T. Press.
- Ullman, S. (1980) Against direct perception. Behavioural and Brain Sciences, **3**, 373-415.
- Ullman, S. (1984) Maximizing rigidity: the incremental recovery of 3-D structure from rigid and non-rigid motion. Perception, **13**, 255-274.
- Ullman, S. (1986) Artificial intelligence and the brain: computational studies of the visual system. Annual Review of Neuroscience, **9**, 1-27.
- Van Santen, J. P. H. & G. Sperling (1985) Elaborated Reichardt detectors. Journal of the Optical Society of America, **A2**, 300-321.
- Verri, A. & T. Poggio (1986) Motion field and optical flow: qualitative properties. M.I.T. Artificial Intelligence Lab. Memo 917 (December, 1986).
- Wallach, H. & D. N. O'Connell (1953) The kinetic depth effect. Journal of Experimental Psychology, **45**, 205-217.
- Warren, W. H. (1984) Perceiving affordances: visual guidance of stair climbing. Journal of Experimental Psychology: Human Perception & Performance, **10**, 683-703.
- Warren, W. H. (1988) Action modes and laws of control for the visual control of movement. In O. Meijer & K. Roth (eds.) Complex Movement Behaviour: 'the' Motor-Action controversy. Amsterdam: North Holland.
- Warren, W. H. & D. J. Hannon (1988) Direction of heading is perceived from optical flow. Nature, **336**, 162-163.

- Warren, W. H. & D. J. Hannon (1990) Eye movements and optical flow. Journal of the Optical Society of America, **A17**, 160-170.
- Warren, W. H. & J. A. S. Kelso (1985) Work group on perception and action. In W. Warren & R. Shaw (ed.), Persistence and Change. Hillsdale, NJ: Lawrence Erlbaum.
- Warren, W. H., J. Morris & M. Kalish (1988) The perception of translational heading from optical flow. Journal of Experimental Psychology: Human Perception & Performance, **14**, 646-660.
- Warren, W. H., Young, D. & D. N. Lee (1986) Visual control of step length during running over irregular terrain. Journal of Experimental Psychology: Human Perception & Performance, **12**, 259-266.
- Waxman, A. M. & S. Ullman (1985) Surface structure and three-dimensional motion from image flow kinematics. International Journal of Robotics Research, **4**, 95-108.
- Waxman, A. M. & K. Wohn (1988) Image flow theory: a framework for 3-D inference from time-varying imagery. In C. M. Brown (ed.), Advances in Computer Vision: vol 1. Hillsdale, NJ.: Lawrence Erlbaum.
- White, B. & G. Mueser (1960) Accuracy in reconstructing the arrangement of elements generating kinetic depth displays. Journal of Experimental Psychology, **60**, 1-11.
- White, B. W., F. Saunders, L. Scadden, P. Bach-y-Rita & C. Collins (1971) Seeing with the skin. Perception and Psychophysics, **7**, 23-27.
- Whittaker, E. T. (1944) A Treatise on the Analytical Dynamics of Particles and Rigid Bodies. New York: Dover.
- Wilkins, B. M. (1970) Analogue and Iterative Methods. London: Chapman & Hall.
- Wood, R. W. (1895) The "haunted-swing" illusion. Psychological Review, **2**, 277-278.
- Woodhouse, N. M. J. (1987) Introduction to Analytical Dynamics. Oxford: O.U.P.

PUBLICATIONS

Perceptual information for the timing of interceptive action

James R Tresilian

Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland
Received 31 March 1989, in revised form 6 February 1990

Abstract. Time-to-contact is an important quantity for controlling activities which involve the timing of interactions with objects and surfaces in motion relative to an observer. Two alternative means for obtaining perceptual information that might be used to obtain the time-to-contact required to correctly time an interaction have been contrasted: a method based on the perception of distance and velocity, and a method due to Lee involving a perceptual variable called tau. A monocular version of the first method is presented and shown to place a highly unrealistic and arbitrary limitation on the capabilities of the visual system. The second method is reviewed and its limitations discussed. Several means by which these limitations can be overcome are presented. Recently reported results from experiments which involved catching self-luminous balls in the dark are interpreted in terms of timing information available to the subject, and the notions of intermodal and multimodal timing information are introduced. Finally, the possibility that timing information is available to an observer which does not involve the variable tau is considered. It is concluded that many questions regarding the perception of time-to-contact remain unresolved and that much empirical research remains to be done.

1 Time-to-contact

Many activities involve precise timing of an interaction with an object or surface in motion relative to the performer. Examples include catching a moving object and hitting an approaching ball with a racquet as in tennis or squash. In one-handed ball catching the precision with which the grasping action of the hand must be timed if the ball is to be held was estimated by Alderson et al (1974) to be of the order of ± 15 ms. Temporal precision of less than 10 ms is probably required of slip fielders in cricket (Lee and Young 1985) and of competition-level ski jumpers (Lee et al 1982): subjects in the experiments of McLeod et al (1985) regularly achieved such precision in a hitting task.

It is generally considered that timing of interceptive acts like catching and hitting is based on perceptual information about the time when the moving object will reach the interception point—the time-to-contact (t_c) of the object with this point. Time-to-contact can be defined as follows: if at some instant of time the distance of the moving object from the interception point (a point on the future path of the object) is d and the relative speed of the object and interception point is v (which will be considered constant), then the time-to-contact (constant velocity) at this instant of time is given by

$$t_c = \frac{d}{v}. \quad (1.1)$$

The question now arises: what perceptual information about t_c is available to an animal or person wishing to intercept a moving object? The primary source for such information for a human being and most other animals will be vision. In what follows the perceptual information about time-to-contact will be analyzed. All the information that is described is defined at an instant of time and is continuously available to the perceptual systems. Thus at every instant a quantity is defined which specifies a time-to-contact at that instant. An observer is, in principle, able to compute this

quantity and use it to control the timing of his movements (subject to a certain perceptuo-motor delay).

In this paper, only the constant velocity case is considered; this greatly simplifies the analysis and may be sufficient when considering the control of interceptive actions since evidence suggests that accelerations are not taken into account in timing tasks (Lee and Reddish 1981; Lee et al 1983; Todd 1981) and that the human visual system displays a lack of sensitivity to acceleration (Gottsdanker et al 1961; Nakayama 1985; Runeson 1974, 1975). However, it must be borne in mind that it is by no means firmly established that human (or animal) performers act according to such a constant velocity approximation.

One answer to the question of what visual information about t_c is available is suggested immediately by equation (1.1): obtain t_c from information about the relevant distance and velocity through the operation of division. This possibility is much discussed in the literature (Cavallo and Laurent 1988; Cavallo et al 1986; Lee 1980a; Lee and Young 1985; McLeod and Ross 1983; Schiff and Detwiler 1979). A second answer has been provided by Lee (1976, 1980a, 1980b) who has shown that it is not necessary for a monocular observer to perceive distance and velocity in order to obtain t_c information. In what follows both these strategies are examined in detail for the case when the relative velocity between object and observer is treated as constant.

2 Timing information during locomotion in a rigid environment

The primary visual input (that which stimulates the receptors) during motion of an observer relative to a rigid environment is a constantly changing or flowing pattern of light intensity imaged on the retinae of the eyes. In order to analyze what visual information about the environment and a perceiver's relation to it is available in such stimulation, a velocity field representation of the visual input is usually employed (eg, Clocksin 1980; Gibson et al 1955; Lee 1974; Longuet-Higgins and Prazdny 1980; Nakayama and Loomis 1974; Waxman and Wohn 1988). An instantaneous velocity field is a projection onto a two dimensional surface (the imaging surface) of the (instantaneous) velocities of the geometrical points on visible environmental surfaces (see below). As such it is an abstraction from the spatio-temporal variation in light intensity constituting the actual stimulation. If information which is found to be present in the velocity field is to be available to a perceiver it must be assumed that the properties of the velocity field defining this information can be extracted from the visual stimulus. It will be assumed that the relevant properties can be extracted; discussion of the problems and assumptions can be found, for example, in Hildreth (1984), Horn (1986), Longuet-Higgins (1986), and Todd (1985).

Bearing the above discussion in mind we can proceed to analyze the monocular *image velocity field*. This will be defined with the use of Longuet-Higgins and Prazdny's camera model of the imaging situation (figure 1). Any point P, with coordinates (X, Y, Z) , on an environmental surface has an instantaneous velocity $\dot{P} = (\dot{X}, \dot{Y}, \dot{Z})$ relative to the camera coordinate system $OX'Y'Z'$. P projects to the point $p = (x, y)$ on the image plane and the instantaneous image velocity $\dot{p} = (\dot{x}, \dot{y})$ is the projection of \dot{P} onto the image plane. The image velocity field is the projection of the velocities \dot{P} of all points P on visible environmental surfaces. Longuet-Higgins and Prazdny (1980) show that the image velocity field is the vector sum of two components: a component due to rotation of the $OX'Y'Z'$ coordinate system relative to the environment, and a component due to translation of this coordinate system.

Only the component due to translation contains any information about the environment or the observer's relation to it. In order to obtain this information from a

general image velocity field, the rotational and translational components need to be separated out. Methods for doing this have been extensively investigated (eg Longuet-Higgins and Prazdny 1980; Prazdny 1980, 1981; Lawton et al 1987).

The translational component or field is a function $v(x, y)$ which assigns to each point (x, y) on the image plane a velocity (\dot{x}, \dot{y}) according to the 'rule' given by

$$\dot{x} = \frac{xV_z - V_x}{Z}, \quad \dot{y} = \frac{yV_z - V_y}{Z}, \tag{2.1}$$

where a dot over a variable represents differentiation with respect to time, a notation adopted throughout. Equation (2.1) can be derived relatively simply from the situation illustrated in figure 1 and is done, for example, in Longuet-Higgins and Prazdny (1980). The velocity vectors of this field are everywhere directed away from or towards a single image point (Longuet-Higgins and Prazdny 1980). All vectors of the field appear to 'flow' out of or into this point which Gibson called the *focus of expansion* (f.o.e.) (Gibson 1950; Gibson et al 1955). The retinal coordinates of the f.o.e. (which will be denoted by $[x_f, y_f]$) depend on the direction of the translational velocity and can be found by noting that at the f.o.e. $(\dot{x}, \dot{y}) = (0, 0)$; substituting this into equation (2.1) yields

$$(x_f, y_f) = \left(\frac{V_x}{V_z}, \frac{V_y}{V_z} \right). \tag{2.2}$$

From (2.1) and (2.2) the following is obtained by direct substitution (Longuet-Higgins and Prazdny 1980),

$$\frac{Z}{V_z} = \frac{x - x_f}{\dot{x}} = \frac{y - y_f}{\dot{y}}. \tag{2.3}$$

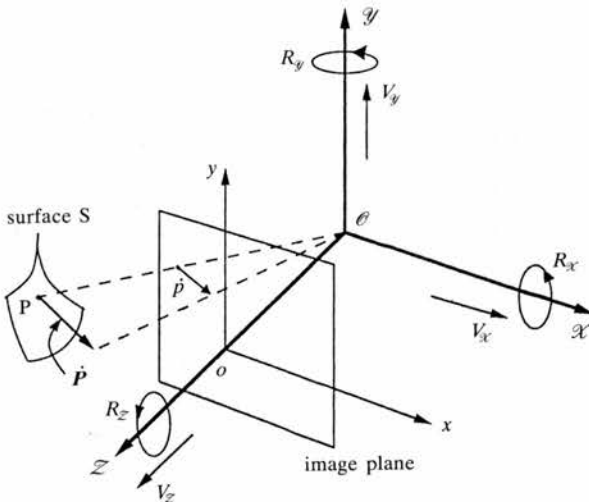


Figure 1. 'Camera' model imaging system, after Longuet-Higgins and Prazdny (1980). θ is the centre of projection (or point of observation) which forms the origin of the cartesian coordinate system $\theta x y z$, with z being the direction of view. The image plane is fixed at unit distance along the z axis in front of the point of projection (purely for convenience) and is oriented parallel to the x - y plane; oxy is the coordinate system of the image plane. Instantaneous rigid body motion of the imaging system relative to the rigid environment (the small surface patch, S , is part of this environment) is represented as an instantaneous translation of the system, (V_x, V_y, V_z) , and an instantaneous rotation about the point of observation, (R_x, R_y, R_z) .

Equation (2.3) demonstrates that the depth (Z) of a point on an environmental surface scaled by the velocity in the Z -direction is specified by variables of the projection (called *optic variables* by, eg, Lee 1980a and Todd 1981). A perceiver can thus obtain the *relative* depths of all points on the visible surfaces in the environment, but not their absolute depths. This is only to be expected since the concept of absolute depth implies the existence of some (absolute) scale of measurement, ie some fixed ('standard') distance in terms of which any depth can be represented. If some known fixed distance is visually available, it is possible to scale the relative depth information; an animal's eye height has been suggested as a possible biological standard (Lee 1974, 1980a; Sedgwick 1973).

Even without a standard distance for scaling depth, the velocity-scaled depth (Z/V_Z) is still informative since it has the dimensions of time and represents the depth in temporal terms. The representation of the translational field given by equation (2.3) can be thought of as defining an instantaneous temporal depth map of the scene in which points in the scene are represented not in terms of their distances in space but in terms of their distances in time. Lee (1976, 1980a, 1980b) stressed that if the Z component of the translational velocity is constant, then the velocity scaled depth is the time remaining before an environmental point reaches the X - Y plane—the *time-to-contact* with this plane. Thus, an animal locomoting through a rigid environment has access to information at every 'instant' (up to a visuo-motor delay) about the t_c with the X - Y plane through the centre of projection of its eye. Of course, in order to use this information appropriately, an observer has also to be able to determine in some way the position and orientation of the X - Y plane. Lee (1976, 1980a, 1980b) has called the velocity field quantity that specifies t_c *tau* (τ). Tau is thus simply the right-hand side of equation (2.3).

The monocular t_c information that has been described does not require that the distance (Z) of an environmental surface and the velocity along the Z -direction be perceived separately and then 'divided' to yield Z/V_Z , the time-to-contact. Fortunately this computation has 'already been done' as it were since Z/V_Z is already expressible in terms of observable quantities of the velocity field [equation (2.3)]. In fact, no one has produced a fully worked out scheme for the perception of t_c based on the division of perceived distance and velocity in the sense of defining the putative distance and velocity information that is used. I shall return to this issue below.

So far the case of t_c information during motion of the observer through a rigid environment has been considered. The definition of this timing information made use of a feature of the global optic flow induced by such motion (the f.o.e.). One is led to ask, therefore, what information is available when no global flow field is defined; for example, when objects are in motion relative to a stationary observer (the definition in terms of the f.o.e. will also not work for objects in motion when the observer is also in motion). This case has been considered by both Lee (Lee and Young 1985) and Todd (1981) whose analyses will be examined in detail in the following sections.

3 Information about moving objects

It will be instructive to consider Todd's (1981) analysis of the situation illustrated in figure 2. The following relations can be derived quite easily (see Todd, 1981):

$$\frac{Z}{A} = \frac{1}{a}, \quad (3.1)$$

$$\frac{\dot{Z}}{A} = -\frac{\dot{a}}{a^2}, \quad (3.2)$$

$$\frac{Y}{A} = \frac{y}{a}, \quad (3.3)$$

$$\frac{\dot{Y}}{A} = \frac{\dot{y}}{a} - \frac{y\dot{a}}{a^2}. \quad (3.4)$$

These equations show that visual information about the (instantaneous) velocity and position of the moving object measured relative to the θyZ coordinate system is available since these quantities scaled in terms of the object size A are given in terms of optic variables. If a perceiver has prior knowledge about the object size, then he or she can obtain useful information about the position (Z, y coordinates) of the object and its velocity. Given equations (3.1) to (3.4) it is a simple matter to show that t_c information is available. Dividing equation (3.1) by equation (3.2) one obtains $-Z/\dot{Z}$, which is the time-to-contact with the y -axis:

$$-\frac{Z}{\dot{Z}} = \frac{a}{\dot{a}} \tag{3.5}$$

The right-hand side of equation (3.5) is a 'retinal' distance divided by a 'retinal' velocity, as are the right hand sides of equation (2.3), and, following Lee, it can be called tau. However, since no global velocity field is used to define tau, here it will be distinguished from tau as defined in the previous section by calling it *local tau* (τ_1).

The above derivation of an expression for t_c can be read in two ways: either as a purely formal argument showing that t_c information is visually available, or as an account of how the visual system might actually compute t_c . If it is read in the latter sense, we get the following account for t_c with the y -axis. The perceiver's visual system computes distance according to equation (3.1) and velocity according to equation (3.2) (it 'knows' about the object's size, A) and then divides these two perceived quantities to obtain t_c . This account imputes a remarkable stupidity and arbitrariness to the perceiver's visual system: it is able to obtain distance and velocity information from equations (3.1) and (3.2) 'directly' as it were, but is unable to exploit the t_c information described by equation (3.5) in the same way, even though (3.5) does not require knowledge of the object size. t_c is rather to be obtained by the curious round-about method of dividing the perceived distance by velocity even though the same result could have been achieved far more easily.

A distance-divided-by-velocity account of the above kind clearly begs the following question: why can the visual system exploit relations that define information about distance and velocity [such as (3.1) and (3.2) above] but not relations that define t_c information? Is it not more reasonable to assume that the visual system with access to basic observable properties of the velocity field (eg a, \dot{a}, y, \dot{y} , and so forth) can

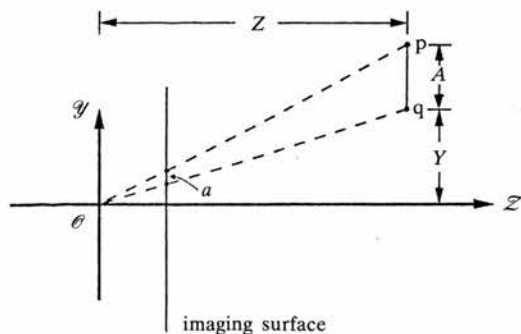


Figure 2. Todd's (1981) model of the imaging system; it can be considered a 'slice' through the imaging system of figure 1. This is a reasonable idealisation since in the real world objects tend to have planar motion. The imaging system is to be considered as fixed relative to the environment and the 'line-segment' \overline{pq} as joining two points p and q on the surface of an object moving through the environment in the y - Z plane. The distance, A , between the two points represents the 'size' of the object in the y dimension. The line segment \overline{pq} is imaged as a line of length a on the imaging surface.

discover relations amongst these observables which define useful information? This is the approach taken in modern computer vision literature on visual motion (see, eg, Waxman and Wohn 1988).

The analysis presented above does show that at least some t_c information about a moving object is available, in principle, to a stationary monocular observer directly in terms of observables of the velocity field. However, this information is very limited indeed; only t_c with the \mathcal{Z} and \mathcal{Y} axes of figure 2 has been shown to be present. If this represents all the t_c information that is available to an observer, it follows that a person will be very constrained in where contact can be made with a moving object and movements still be timed precisely. This prediction does not appear to be borne out in everyday experience; it is commonly felt, for example, that a ball can be caught almost anywhere within reach. Indeed, von Hofsten (1983, 1987) has argued that the catching skills of very young infants are more flexible than would be predicted by the above analysis, leading him to doubt the utility of tau as information for timing interceptive acts like catching (von Hofsten and Lee 1985). In the next section the possibility of more general timing information based on tau is investigated.

4 General time-to-contact information

For the sake of mathematical convenience a spherical projection surface will be employed for the analyses presented in this section (any expressions obtained for a spherical surface can be uniquely transformed into corresponding expressions for a plane surface: see, eg, Lee (1974) and Longuet-Higgins and Prazdny (1980). A simple result, quoted from Lee and Young (1985), will be used here: for a spherically symmetric object approaching a point of observation with constant speed the instantaneous t_c with the observation point is given by $2\Omega/\dot{\Omega}$ (for small Ω), where Ω is the solid angle subtended by the object at the observation point (the simple derivation is given in full in Lee and Young 1985). $2\Omega/\dot{\Omega}$ will be denoted τ_1 (cf. Lee and Young 1985).

Von Hofsten (1983) showed that when an object is moving past an observer (figure 3), if τ_1 as defined above is used as t_c information, significant timing errors will result, which are not observed even in the catching skills of von Hofsten's infant subjects. In response to this Lee and Young (1985; see also von Hofsten and Lee 1982) showed that 'time-to-nearest-approach' information is visually available (ie the time to contact with the point N in figure 3). On denoting the time-to-nearest-approach by t_N Lee and Young's expression for it becomes

$$t_N = \frac{\tau_1}{1 + \tau_1^2 \dot{\theta}^2}. \quad (4.1)$$

To obtain this equation one can differentiate the two relations $\cos \theta = L/R$ and $\sin \theta = D/R$, substitute appropriately, and use the relation $\tau_1 = -R/\dot{R}$. Nevertheless, a perceiver armed with t_N information is still very constrained in where catches can be accurately timed, and in the opinion of von Hofsten (von Hofsten and Lee 1985) this equation does not appear to capture the flexibility of human catching skills. It would be more satisfactory to be able to show that less-constraining timing information is visually available.

Consider again figure 3: suppose some arbitrary point p on the upcoming path of the object represents a chosen point of interception. By the sine rule we can write

$$\sin \beta = \frac{R \sin \psi}{D_p}, \quad (4.2)$$

differentiating this yields

$$\frac{\dot{D}_p}{D_p} \sin \psi = \frac{\dot{R}}{R} \sin \psi + \dot{\psi} \cos \psi, \quad (4.3)$$

since the angle β is constant. Now D_p/\dot{D}_p (where \dot{D}_p is identical to the speed of the object, $|V|$) is the time-to-contact of the object with the point p (which will be denoted t_p) and $-R/\dot{R} = \tau_1$. Substituting t_p and τ_1 into equation (4.3) one obtains, with some rearrangement, the following relation:

$$t_p = \frac{\tau_1}{1 + \tau_1 \psi \cot \psi} \tag{4.4}$$

Equation (4.4) shows that at every instant of time the time remaining before a moving object reaches any point p on its upcoming path is visually available as long as the velocity of the object remains constant. Thus timing information which in principle allows an object to be intercepted successfully almost anywhere within reach is at least potentially available.

To define the information described by (4.4) only the *direction* of the point p is required since this suffices to specify the angle ψ . It would appear, therefore, that the direction of the point p needs to be visually available. It would be available if p were taken to be in the direction of some visible feature on a surface in the fixed environment; this feature then labels the direction of the interception point p. If such an environmental label for the interception point is required then the use of the information described by equation (4.4) would require that the observer has a structured visual environment. Under such conditions an observer wishing to intercept the object is able to do so along any direction from the point of observation through which the object will pass provided that some visible environmental feature lies in that direction. This assumes, of course, that the perceptual source of the variables appearing in equation (4.4) must be vision. In the next section the possibility that they can be obtained with the use of other perceptual systems is examined.

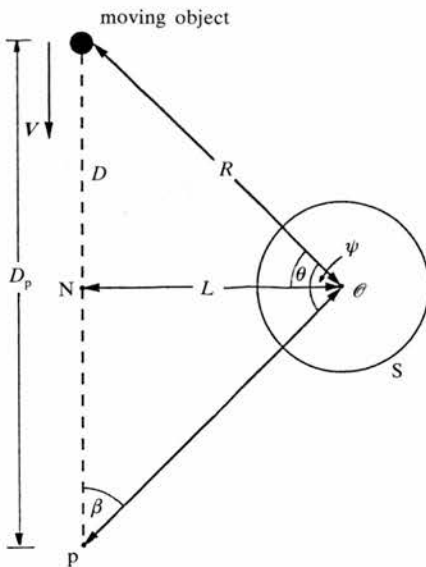


Figure 3. Geometry of time-to-contact. A spherical object (eg, a ball) is moving in a plane; for this reason a two-dimensional model of the imaging situation is employed as in section 3 (Lee and Young 1985; Todd 1981). An object is moving (instantaneously) in the plane of the diagram along the path shown as a dotted line past the point of observation (centre of projection) θ . The point N on the projected path of the object is the point of nearest approach to the observation point, were the object to continue with the velocity V . The circle S represents a slice through a spherical imaging surface centred on θ .

5 Intermodal and multimodal timing information

Recently some experiments have been conducted where the task consisted of catching self-luminous balls in an otherwise completely dark room (Rosengren et al 1988; Savelsburgh et al 1989). Experimental manipulations were made which involved providing subjects with selected visual information about themselves (eg the catching hand) or about their environment by making portions of it self-luminous (see below). With only the ball visible, subjects were found to be quite capable of effecting clean catches, though considerably less often than under conditions of full room illumination. How is this catching ability to be understood when the variables ψ , $\dot{\psi}$, and $\dot{\theta}$ appearing in equations (4.1) and (4.4) are not visually available?

Consider the variable $\dot{\theta}$ which appears in equation (4.1) describing time-to-nearest-approach information (Lee and Young 1985). As pointed out by Lee and Young this is the rate of change of direction of the object relative to the point of observation. Likewise, the variable $\dot{\psi}$ appearing in equation (4.4) is this rate of change of direction (the rate of change of direction will be denoted u in what follows). Lee and Young (1985) observed that at a stationary point of observation this quantity is equal to,

“the rate of change of the angle subtended at the eye by the object and any fixed direction which lies in the plane containing the eye, the object and the relative velocity vector.”
(Lee and Young 1985, page 5)

A person wishing to catch an object will typically turn his or her head and generate pursuit eye-movements which act to keep the image of the object on the foveal region of the retinae. This being the case, the rate of change of direction of the object is equal to the rate of rotation of the eyes relative to the environment (if tracking is perfect). In illuminated environments such head-eye movements generate retinal flows which specify the rotation of the eyes in the environment—flows with a purely propriospecific significance (Koenderink and van Doorn 1981; Longuet-Higgins and Prazdny 1980). In the dark no such flows are generated, of course. Nevertheless, if the head is fixed and only the eyes free to move, the rate of rotation of the eyes in the head as they follow the moving object is equal to its rate of change of direction. In this case the rate of change of direction of the object is potentially available from extravisual sources, since the rate of rotation of the eye in the head could in principle be obtained with the use of proprioceptive/efference copy information. If the head is free to move, then the rate of change of direction of the object is given jointly by the rate of rotation of the head on the shoulders and the rate of rotation of the eyes in the head. (Note that there will be some translation of the observation point when the head rotates and indeed when the eye rotates. It will be ignored here for the sake of illustration). If the rate of change of direction of the object can be supplied by perceptual systems other than vision, then the timing information described by equations (4.1) and (4.4) is potentially *multimodal* timing information. Here ‘multimodal’ means that the observables appearing in equations describing perceptual information come from more than one perceptual source.

Consider now the angular variable ψ which appears in equation (4.4). This quantity is the angle between the (instantaneous) direction of the moving object and the direction of the interception point (measured at the observation point). As described above, ψ is potentially available visually in an illuminated environment; when catching luminous balls in the dark, however, it is not available from this perceptual source. If it is supposed that the direction of the interception point is available from some extravisual source, then the angle ψ is, at least in principle, available to an observer, since both directions needed to define ψ are perceptually specified: the direction of the ball is visually specified, the direction of the intercept point is specified extravisually. It thus requires two sources of information to define ψ in this

case; if both sources are perceptual, then ψ is defined across perceptual systems as it were. In this case it seems appropriate to say that ψ is *intermodally* specified (it is defined as a relationship *between* perceptual systems).

These notions of multimodal and intermodal information may help to explain the ability to time correctly catches of self-luminous balls in the dark since they allow the variables ψ and u to be perceptually available in the absence of visual cues. However, it may be the case that vision is able to supply these quantities with more precision than proprioceptive/efference copy sources. There is evidence that vision is the most precise source of information about movement of the head relative to a fixed environment (see, eg, Johansson 1977; Lee 1978). If vision is indeed more precise, then supplying a person with vision of the environment sufficient to provide the rate of change of direction (u) of a moving object should improve the accuracy with which a catch can be timed. This may be the explanation of results reported by Rosengren et al (1988) and Savelsburgh et al (1989). Rosengren et al reported that when self-luminous strips were attached to the wall of the experimental room behind the point of ball projection, catching performance was significantly better than performance when only the ball was visible. Savelsburgh et al reported a similar result: they found that a self-luminous grid, placed such that the (luminous) ball was seen moving in front of it, significantly improved the accuracy of the timing of catching relative to conditions in which only the ball was seen.

The argument presented in this section can be put into a more general form as follows. Intuitively, perceptual information about the value of some environmental property or relation between an observer and his environment, E , may be conceived of as a functional relationship of the form

$$E = f(\pi_1, \pi_2, \dots, \pi_n), \quad (5.1)$$

where the n quantities π are perceptual variables, ie variables available directly to the perceptual systems. Visual information is then a relation where all the variables π are 'optic' variables; multimodal information is a relation where the variables π do not all have their source in the same perceptual system; and intermodal information is a relation where some of the variables π are defined only between perceptual systems. Note that an intermodal relation may be expressed as a purely multimodal relation by suitably redefining the variables π . Whether information is to be considered intermodal or not depends on what variables are available to the perceptual systems.

In information science, information is something abstract and structural; it does not depend on the physical medium that carries it. It may thus be carried in many different media and transduced from one medium to another. Information in this sense, as pattern or structure, is that of Shannon's information theory (Shannon and Weaver 1949). Hence information itself is intrinsically *amodal* (ie is independent of the sensory modalities which carry it, cf Gibson 1966). Perceptual information can then be considered as being obtained either unimodally, multimodally, or intermodally. This notion that information is abstract is really very simple and commonplace: for example, the property of being square in shape is abstract in this sense, since it does not depend on the nature of the physical substance from which the square is made, the physical medium which realises it.

6 Limitations of the account

In section 4 the quantity τ_1 was defined to be $2\Omega/\dot{\Omega}$ which can be considered as the inverse of the rate of dilation of the image of a moving object on a spherical projection surface. In section 4 the moving object was considered to be spherical (eg, a ball). There is good reason for this; it guarantees that any changes in the size of the

image of the object are due to changes in the distance of the object from the observation point. When the object is not spherical, there can be changes in the image size that are due to different regions of the surface of the object being visible as the direction of the object from the imaging system changes as it moves. If the object was stationary, similar changes in its image would be generated by rotating the object (clearly, rotating a spherical object does not change the size of its image). The same is true for a planar imaging surface. In Todd's situation (figure 2) the 'line segment' lies parallel to the image plane; it is this fact that guarantees that changes in the image size are due solely to changes in the distance of the line segment from the image plane.

To examine this situation in more detail we will return to Longuet-Higgins and Prazdny's imaging system, figure 1. Now the relative translational velocity between the surface and the imaging system can be taken as being due to motion of the surface in the environment, the imaging system itself being considered stationary. The surface will be taken to be lying instantaneously along the Z -axis (the 'line-of-sight') and as having a slope α in the X direction and β in the Y direction in the neighbourhood of the point where it is pierced by the Z -axis (following the notation used by Longuet-Higgins and Prazdny). A measure of the rate of dilation of the local velocity field (at the origin of the image plane coordinate system) generated by projection of the relative motion of the imaging system and the moving surface onto the image plane is provided by the divergence of the velocity field at the origin. If the velocity field $v(x, y)$ has an x component v_x and a y component v_y , then the divergence (μ) is given by,

$$\mu = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}. \quad (6.1)$$

Longuet-Higgins and Prazdny (1980) derive expressions for the spatial derivatives appearing on the right-hand side of this expression. Since the derivation of these expressions does not directly concern us here, Longuet-Higgins and Prazdny's results will simply be quoted (see also Koenderink and van Doorn 1975). They are, respectively,

$$\frac{\partial v_x}{\partial x} = \frac{\alpha V_x}{Z} + \frac{V_z}{Z} \quad \text{and} \quad \frac{\partial v_y}{\partial y} = \frac{\beta V_y}{Z} + \frac{V_z}{Z}. \quad (6.2)$$

Clearly, therefore, the divergence is given by

$$\mu = \frac{\alpha V_x + \beta V_y}{Z} + \frac{2V_z}{Z}. \quad (6.3)$$

If the first term of (6.3) is zero, either because α and β are both zero (the surface is not sloped) or because both V_x and V_y are zero (the object is on a direct collision course with the observation point), we obtain

$$\frac{2}{\mu} = \frac{Z}{V_z}. \quad (6.4)$$

That is, twice the inverse of the divergence is equal to the time-to-contact. This matches Lee and Young's result given above, where the time-to-contact (Z/V_z) is given by $2\Omega/\dot{\Omega}(\tau_1)$ which is numerically equal to twice the inverse of the divergence (Koenderink 1985). However, if the slopes are not both zero and the object is not on a collision course with the observation point, then contributions to the value of the divergence due to the first term in equation (6.3) mean that the inverse of the divergence does not give information about the time-to-contact.

The analysis presented in the present section would appear to seriously compromise the generality of the analysis provided in section 4. However, owing to the fact that in most situations motion of the objects which one might wish to intercept (in catching or hitting, for example) will be confined to lie on a plane, the generality of analysis is not significantly compromised. In figure 2, the motion of the object, idealised as the 'line segment' \overline{pq} , is considered to lie in the plane of the page. If this object has a component of velocity parallel to the image plane, then the inverse of the rate of dilation of \overline{pq} will give t_c information only when the two points on the surface of the object used to define \overline{pq} are such that it is parallel to the image plane. If \overline{pq} is sloped relative to the image plane, then we cannot use its rate of dilation as t_c information, since there will be a contribution which depends on the slope and the velocity parallel to the image plane. Nevertheless, since the object in figure 2 moves in the plane of the page, it has no component of velocity perpendicular to the page—the \mathcal{X} direction ($V_x = 0$). Thus, the inverse of the rate of dilation measured along the \mathcal{X} direction is equal to the t_c . This can be seen by considering the relations in (6.2): the 'divergence' along the \mathcal{X} direction alone is equal to $\partial v_x / \partial x = \alpha V_x / Z + V_z / Z$, the first expression in (6.2). Since the velocity V_x is zero the first term vanishes leaving V_z / Z which is t_c^{-1} .

It can be concluded, therefore, that, in catching and hitting irregularly shaped objects, t_c information can be obtained by measuring the inverse of the local rate of dilation (τ_1) along the dimension perpendicular to the plane of motion of the object. In the next section the possibility that τ_1 can be dispensed with entirely in the timing of interceptive actions is considered.

7 Timing without tau

Two possible means by which the use of local tau can be avoided are presented in this section. First, an account of catching objects moving along linear paths proposed by von Hofsten is examined and, second, a method for avoiding use of tau by exploiting binocular information is presented.

7.1 Von Hofsten's missile strategy

Consider the following situation. An object is moving along a straight horizontal path through the air at constant speed. A missile launcher below it on the ground is faced with the task of shooting down the object. If the launcher moves along the ground with the same horizontal velocity as the object, then there will be no motion of the object relative to the launcher. A missile need only be directed at the object and fired for the object to be shot down. If the launcher is fixed to the ground, then the same result could be achieved by calculating the speed and direction of the object relative to the ground and firing a missile such that it has this velocity combined with another component of velocity directed at the instantaneous position of the object. The missile is thus fired with a velocity which is the vector sum of these two components and it will hit the object at a point C along its path (figure 4). This is essentially the strategy for object interception that von Hofsten has proposed as a model of catching (von Hofsten 1983, 1987), though he included the feature of continuous feedback regulation (the guided missile strategy) and recognized that, for a person whose position is fixed, only part of the path of the object will be reachable.

In the ballistic case of von Hofsten's strategy (the argument to be developed is unaffected by whether or not feedback regulation is included), a person wishing to catch an object must select an interception point somewhere on the reachable portion of the trajectory. Intercepting the object then amounts to finding the velocity in the direction of the object which, when added to its perceived velocity (as in figure 4), gives a resultant velocity in the direction of the interception point.

At first glance this strategy appears to be an attractive alternative to a scheme based on the t_c information described earlier because the catcher can apparently both position and time the catch using only perceived velocity information. There are, however, two problems with this strategy. First, it fails to explain how the grasp of a catch can be timed—the grasp still appears to require t_c information. Second, it makes the tacit assumption that the velocity of the hand is constant. However, this velocity will not be constant; a person reaching out to make a catch starts the hand from rest and brings it to rest again at, or shortly after, contact with the object (cf Alderson et al 1974). This has the consequence that t_c information is required to achieve interception when the missile strategy is used. The reason is that, in order for interception to be achieved, the average speed of the hand as it travels from its initial position to the point of interception must be equal to the magnitude of the velocity vector calculated for that intercept (figure 2). Thus the catcher must execute a movement of the hand of this average speed; this cannot be done without knowledge of the distance to be travelled. Once this is known, the catcher simply reaches the distance in a time, t , given by

$$t = \frac{\text{distance to be reached}}{\text{average speed}}. \quad (7.1)$$

This time is simply the time-to-contact of the object with the interception point; exactly the information von Hofsten argued was not required. Thus, the missile strategy when applied to the catching task that von Hofsten discussed is a restatement of the idea that t_c information is obtained from distance and velocity through division. How the required distance and velocity information might be obtained has not been discussed. Nonetheless, if such information is perceptually available, it is possible not only to time the catch but to position the hand correctly as well (at least in the case when the path of motion of the object is a straight line).

That the required velocity information is available is indicated by, for example, the analysis of Todd (1981) discussed earlier. The problem of obtaining the distance of the interception point remains. That this information is available follows from the above discussion. If a direction in which the object is to be intercepted is chosen, then an average velocity for the hand could be calculated according to von Hofsten's model (figure 4). As shown in the last section, the time-to-contact of the object with the interception point is perceptually available when only the direction of the interception point is known [equation (4.4)]. Thus both the time t and the 'average velocity' appearing in equation (7.1) can be obtained, and hence one can derive the distance to be reached.

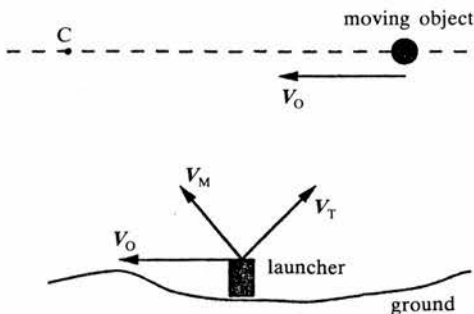


Figure 4. The 'missile strategy'. A missile launcher (shown as a textured block) is on the ground below an object moving along a linear path shown as a dotted line. A missile fired with a velocity V_M —the sum of a component V_T towards the instantaneous position of the moving object, and a component V_O equal to the velocity of the moving object—will intercept the object at a point C along its path.

7.2 t_c information without tau

Consider either a moving object heading for a collision with some other object or surface a long way from the observer or a 'ball' moving towards a 'bat' in a game of video tennis. In both cases the two 'objects' in collision are known to lie in the same plane (in the first case the great distance guarantees that even if they are not coplanar it is likely to be a good approximation to assume that they are). Coplanarity means that we can apply the optic geometry diagrammed in figure 5. From this geometry it is simple to derive the following relation (see caption to figure 5) which shows how the time remaining before the two objects collide is visually available:

$$t_c = \frac{w}{\dot{w}} \tag{7.2}$$

This is rather obvious, but can be used in the derivation of a more general relation giving potentially the same information as equation (4.4).

Considering the geometry diagrammed in figure 6, one can obtain the following relation by application of the sine rule,

$$D = W \frac{\sin \varphi}{\sin \vartheta} \tag{7.3}$$

Differentiating this, one obtains

$$V = \frac{\sin \vartheta (\dot{W} \sin \varphi + W \dot{\varphi} \cos \varphi) - W \dot{\vartheta} \sin \varphi \cos \vartheta}{\sin^2 \vartheta} \tag{7.4}$$

where V is the speed of the object. Dividing (7.4) by (7.3) yields

$$\frac{V}{D} = \frac{\dot{W}}{W} + \dot{\varphi} \cot \varphi - \dot{\vartheta} \cot \vartheta \tag{7.5}$$

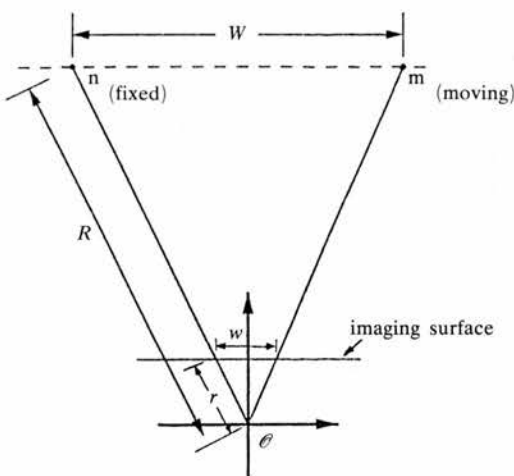


Figure 5. Optic geometry of time-to-contact in the video-game case. Two collinear points m and n project onto a slice through a plane imaging surface. If the point n is considered motionless and the other, m , as moving towards it in the plane of the diagram, then by similar triangles we have $W/R = w/r$ which when differentiated with respect to time gives $\dot{W}/R = \dot{w}/r$. Division of these two relations yields equation (7.2) given in the text (where $t_c = W/\dot{W}$). Note that this argument is unchanged when points m and n are both moving, as is readily verified.

Noting that V/D is the reciprocal of the time-to-contact t_p , that $\dot{W}/W = \dot{w}/w$ and $\vartheta = 180 - \varphi - \beta$, where b is constant, one obtains from (7.5):

$$t_p^{-1} = \frac{\dot{w}}{w} + \dot{\varphi} [\cot \varphi - \cot(\varphi + \beta)]. \quad (7.6)$$

All the variables appearing on the right-hand side of equation (7.6) are potentially available from the projection except the angle (β) that the velocity vector makes with the image plane. How can this angle be obtained? Lee and Young (1985) discuss a monocular method for obtaining the direction of the velocity vector which involves the use of τ_1 [see also Todd (1981) for a similar method]. If such a method is used, then equation (7.6) has no advantages over equation (4.4) and is considerably more long-winded. It is possible, however, that a binocular observer can obtain the direction of the velocity in a manner independent of τ_1 , by making use of the binocular flow information about the direction of motion in depth described by Regan and Beverley (1979).

That the information is binocularly available can be demonstrated by introducing a second imaging system into figure 6 slightly to the left, say, of the system illustrated, and having its image plane oriented parallel to that of the illustrated system. We can clearly write the following two equations for these two imaging systems, indicating that variables are defined for the left-hand system by the subscript L and for the right-hand system by the subscript R;

$$\frac{1}{t_p} = \frac{\dot{w}_R}{w_R} + \dot{\varphi}_R [\cot \varphi_R - \cot(\varphi_R + \beta)]$$

and

$$\frac{1}{t_p} = \frac{\dot{w}_L}{w_L} + \dot{\varphi}_L [\cot \varphi_L - \cot(\varphi_L + \beta)].$$

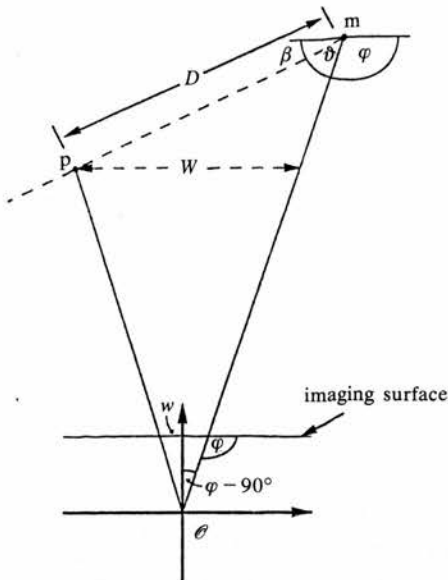


Figure 6. Instantaneous optic geometry of time-to-contact for arbitrary planar motion. A moving point m has a velocity V in the direction of a point p making angle β to the imaging surface. The angular quantities can be transformed into variables of planar projection but the corresponding expressions for the time-to-contact information are more complicated, and for the sake of clarity angular quantities are represented as such in the text.

We now have two equations in two unknowns (t_p and β) that can be solved for t_p , which is thus visually specified without requiring local tau (τ_l). Whether or not the information described here is actually used by an observer is a matter for empirical investigation.

8 Conclusions

The following conclusions may be drawn from the discussion presented in this paper:

(i) The 'optic' variable tau is potentially useful in providing timing information either defined 'globally', through the f.o.e. of the translational component of the image velocity field generated by motion through the environment, or defined locally as the inverse of the rate of dilation of the image of an object. However, the skilled interceptive actions of humans (and probably other animals) almost certainly require timing information that tau alone does not appear to provide. More general timing information involving perceptual variables other than tau is, in principle, available to an observer (section 4). This suggests that there is more to interceptive timing than tau. Empirical research is required to establish if the other variables described in this paper are important in the perception of time-to-contact.

(ii) The timing information described in this paper does not necessarily represent purely visual information, since certain variables appearing in the formal definition of this information may be obtained from perceptual sources other than vision.

(iii) The strategy for obtaining time-to-contact information from previously obtained distance and velocity information attributes to the visual system a peculiar limitation which is difficult, if not impossible, to justify (at least in the monocular case considered in section 3). If a version of this strategy is to be presented as a serious account of how timing information might be obtained, it needs to be shown how the attribution of such a limitation can be avoided (or, if possible, justified). It is curious that the distance-divided-by-velocity strategy is so often discussed in the literature, considering that it has never been properly worked out. It is perhaps because such an approach is deemed to be 'computational' (see, eg, McLeod and Ross 1983), whereas Lee's approach is often taken as the paradigm example of Gibson's 'direct' perception (see, eg, Turvey and Carello 1986; Turvey et al 1981). However, all that 'directness' amounts to in this case is that t_c can be expressed as a relation between observables of the optic projection. Much of the work in contemporary computer vision analyzing the information in 'optic flows' has exactly this character: it involves showing how environmental properties and observer environment relations can be expressed in terms of relations between observables of the optic projection (see, eg, Clocksin 1980; Koenderink 1985; Lawton et al 1987; Longuet-Higgins and Prazdny 1980; Prazdny 1983; Subbarao 1988; Waxman and Wohn 1988). Lee's demonstration of the optical specification of time-to-contact could serve as an example of the kind of analysis required in computer vision.

(iv) The account of object interception described by von Hofsten and purporting to involve only velocity information was shown not to obviate the need for time-to-contact information, contrary to von Hofsten's claim.

(v) The use of local tau as a variable important in defining time-to-contact was shown to involve certain complications due to the fact that size change in the image of an object can occur for reasons other than movement towards or away from an observer, even for a rigid object. Although the complications are not severe, it is possible to obtain information about time-to-contact without the use of tau as was shown in section 7. Whether local tau (inverse of image dilation) is used in interceptive timing under normal viewing conditions therefore remains an open empirical question.

Acknowledgements. I would like to thank Dave Lee for commenting on the work reported in this paper and for much helpful discussion concerning 'time-to-contact'. The work presented here was completed whilst the author was in receipt of a postgraduate studentship awarded by the Medical Research Council of Great Britain.

References

- Alderson G J K, Sully D, Sully H, 1974 "An operational analysis of a one-handed catching task using high speed photography" *Journal of Motor Behavior* **6** 217-226
- Cavallo V, Laurent M, 1988 "Visual information and skill level in time-to-collision estimation" *Perception* **17** 623-632
- Cavallo V, Laya O, Laurent M, 1986 "The estimation of time-to-collision as a function of visual stimulation" in *Vision in Vehicles* Eds A G Gale, M Freeman, C Haselgrave, P Smith, S Taylor (Amsterdam: Elsevier) pp 365-374
- Clocksink W F, 1980 "Perception of surface slant and edge labels from optical flow: a computational approach" *Perception* **9** 253-269
- Gibson J J, 1950 *The Perception of the Visual World* (Boston, MA: Houghton-Mifflin)
- Gibson J J, 1966 *The Senses Considered as Perceptual Systems* (Boston, MA: Houghton-Mifflin)
- Gibson J J, 1979 *The Ecological Approach to Visual Perception* (Boston, MA: Houghton-Mifflin)
- Gibson J J, Olum P, Rosenblatt F, 1955 "Parallax and perspective during aircraft landings" *American Journal of Psychology* **68** 372-385
- Gottsdanker R M, Frick J, Lockerd R, 1961 "Identifying the acceleration of visual targets" *British Journal of Psychology* **52** 31-42
- Hildreth E C, 1984 *The Measurement of Visual Motion* (Cambridge, MA: MIT Press)
- Hofsten C von, 1983 "Catching skills in infancy" *Journal of Experimental Psychology: Human Perception and Performance* **9** 75-85
- Hofsten C von, 1987 "Catching" in *Perspectives on Perception and Action* Eds H Heuer, A F Sanders (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 95-111
- Hofsten C von, Lee D N, 1985 "Dialogue on perception and action" in *Persistence and Change* Eds W H Warren, R E Shaw (Hillsdale, NJ: Lawrence Erlbaum Associates)
- Horn B K P, 1986 *Robot Vision* (Cambridge, MA: MIT Press) pp 231-242
- Johansson G, 1975 "Studies on the visual perception of locomotion" *Perception* **6** 365-376
- Koenderink J J, 1985 "Space, form and deformations" in *Brain Mechanisms and Spatial Vision* Eds D Ingle, M Jeannerod, D N Lee (Dordrecht, Netherlands: Martinus Nijhoff) pp 31-58
- Koenderink J J, Doorn A J van, 1975 "Invariant properties of the motion parallax field due to movement of rigid bodies relative to an observer" *Optica Acta* **22** 773-791
- Koenderink J J, Doorn A J van, 1981 "Exterospesific component of the motion parallax field" *Journal of the Optical Society of America* **71** 953-957
- Lawton D, Rieger J, Steenstrup M, 1987 "Computational aspects of visual motion processing" in *Vision, Brain and Cooperative Computation* Eds M Arbib, R Hanson (Cambridge, MA: MIT Press)
- Lee D N, 1974 "Visual information during locomotion" in *Perception: Essays in Honour of J.J. Gibson* Eds R Macleod, H Pick (Ithaca, NY: Cornell University Press) pp 250-267
- Lee D N, 1976 "A theory of the visual control of braking based on time-to-collision" *Perception* **5** 437-459
- Lee D N, 1978 "The functions of vision" in *Modes of Perceiving and Processing Information* Eds H Pick, E Saltzman (Hillsdale, NJ: Lawrence Erlbaum Associates)
- Lee D N, 1980a "The optic flow field: the foundation of vision" *Philosophical Transactions of the Royal Society of London B* **290** 169-179
- Lee D N, 1980b "Visuo-motor coordination in space-time" in *Tutorials in Motor Behaviour* Eds G E Stelmach, J Requin (Amsterdam: North Holland) pp 281-295
- Lee D N, Lishman J, Thompson J A, 1982 "Regulation of gait in long jumping" *Journal of Experimental Psychology: Human Perception and Performance* **8** 448-459
- Lee D N, Reddish P E, 1981 "Diving gannets: a paradigm of ecological optics" *Nature (London)* **293** 293-294
- Lee D N, Young D S, 1985 "Visual timing of interceptive action" in *Brain Mechanisms and Spatial Vision* Eds D Ingle, M Jeannerod, D N Lee (Dordrecht, Netherlands: Martinus Nijhoff) pp 1-30
- Lee D N, Young D, Reddish P, Lough S, Clayton T, 1983 "Visual timing in hitting an accelerating ball" *Quarterly Journal of Experimental Psychology A* **35** 333-346
- Longuet-Higgins H C, 1986 "Visual motion ambiguity" *Vision Research* **26** 181-183

-
- Longuet-Higgins H C, Prazdny K, 1980 "The interpretation of moving retinal image" *Proceedings of the Royal Society of London B* **208** 385-397
- McLeod P, McLaughlin C, Nimmo-Smith P, 1985 "Information encapsulation and automaticity. Evidence from the visual control of finely timed actions" in *Attention & Performance XI* Eds M Posner, O Marin (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 391-406
- McLeod R W, Ross H E, 1983 "Optic flow and cognitive factors in time-to-collision estimates" *Perception* **12** 417-423
- Nakayama K, 1985 "Biological image motion processing: a review" *Vision Research* **25** 625-660
- Nakayama K, Loomis J M, 1974 "Optical velocity patterns, velocity sensitive neurons and space perception: a hypothesis" *Perception* **3** 63-80
- Prazdny K, 1980 "Egomotion and relative depth map from optical flow" *Biological Cybernetics* **36** 87-102
- Prazdny K, 1981 "Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer" *Computer Graphics and Image Processing* **17** 238-248
- Regan D, Beverley K, 1979 "Binocular and monocular stimuli for motion in depth. Changing-disparity and changing-size feed the same motion-in-depth stage" *Vision Research* **19** 1331-1342
- Rosengren K, Pick H, Hofsten C von, 1988 "The role of visual information in ball catching" *Journal of Motor Behavior* **20** 150-164
- Runeson S, 1974 "Constant velocity—not perceived as such" *Psychological Research* **37** 3-23
- Runeson S, 1975 "Visual prediction of collision with natural and unnatural motion functions" *Perception & Psychophysics* **18** 261-266
- Savelsburgh G J, Whiting H T A, Bootsma R J, 1989 "An ecological perspective on the control of one-handed catching" in *Proceedings of the International Congress of Psychology, Sydney, Australia*
- Schiff W, Detwiler M L, 1979 "Information used in judging impending collision" *Perception* **8** 647-658
- Sedgwick H, 1973 "The visible horizon: a potential source of visual information for the perception of size and distance" *Dissertation Abstracts International* **34** 1301B-1302B (University Microfilms no. 73-22530)
- Shannon C E, Weaver W, 1949 *Mathematical Theory of Communication* (Urbana, IL: University of Illinois Press)
- Subbarao M, 1988 *Interpretation of Visual Motion: A Computational Study* (London: Pitman)
- Todd J T, 1981 "Visual information about moving objects" *Journal of Experimental Psychology: Human Perception and Performance* **7** 795-810
- Todd J T, 1985 "Perception of structure from motion: is projective correspondence of moving elements a necessary condition?" *Journal of Experimental Psychology: Human Perception and Performance* **11** 689-710
- Turvey M T, Carello C, 1986 "The ecological approach to perceiving-acting" *Acta Psychologica* **63** 133-155
- Turvey M T, Shaw R E, Reed E S, Mace W, 1981 "Ecological laws of perceiving and acting. In reply to Fodor and Pylyshyn" *Cognition* **9** 237-304
- Ullman S, 1979 *The Interpretation of Visual Motion* (Cambridge, MA: MIT Press)
- Waxman A M, Wohn K, 1988 "Image flow theory: a framework for 3-D inference from time-varying imagery" in *Advances in Computer Vision: Volume 1* Ed. C M Brown (Hillsdale, NJ: Lawrence Erlbaum Associates)

Empirical and Theoretical Issues in the Perception of Time to Contact

James R. Tresilian

University of Edinburgh, Edinburgh, Scotland

Four questions concerning the perceptual source of information about time to contact (t_c) are addressed: (a) What conditions are required for the optic variable tau to play a role in the perception of t_c ? (b) When these conditions are met, does tau alone provide sufficient information for accurate timing of interceptive actions? (c) Does a distance divided by velocity account of t_c perception provide a convincing alternative to an account that is based on tau? (d) Is there any empirical evidence that distinguishes the two accounts? A "global" type of tau variable and a "local" type of tau variable are distinguished, each with different limitations. The discussion is largely concerned with local tau variables, 2 versions of which are identified. It is concluded that tau alone cannot provide sufficient information for skilled timing. An extended tau-based account presented in an earlier article (Tresilian, 1990) is discussed. It is argued that no extant empirical data can distinguish the extended account from the distance divided by velocity account.

Over the past two decades there has been considerable interest in what perceptual information enables veridical perceptual estimation of time to contact (Cavallo & Laurent, 1988; Cavallo, Laya, & Laurent, 1986; Groeger & Brown, 1988; Laurent, Dinh Phung & Ripoll, 1989; Lee, 1976, 1980; Lee & Reddish, 1981; Lee & Young, 1985; Lee, Young, Reddish, Lough, & Clayton, 1983; McLeod, McLaughlin, & Nimmo-Smith, 1985; McLeod & Ross, 1983; Schiff & Dettwiler, 1979; Schiff & Oldak, 1990; Solomon, Carello, & Turvey, 1984; Todd, 1981). Time to contact (t_c) is potentially a useful variable for the control of actions such as catching and hitting, which involve interacting with moving objects. Its importance as a control variable has been argued cogently by Lee and others (e.g., Lee, 1976, 1980; Turvey & Carello, 1986).

Lee (1976, 1980) provided a theory of the visual information that might be used in the perception of t_c . This theory centers on Lee's demonstration (following from the earlier work of Carel, 1961; Hoyle, 1961; Gibson, Olum, & Rosenblatt, 1955; Purdy, 1958) that t_c with an observer is optically specified by a single optic variable, which Lee called tau (τ , defined in the following). Perception of t_c through τ has often been presented as the paradigm example of Gibson's "direct perception" (e.g., Michaels & Carello, 1981; Turvey & Kugler, 1984), and perhaps this accounts for the considerable attention it has received. Whatever one's views on the issue of direct versus indirect perception, the most often discussed question in the literature on t_c is a consequence of just this issue: whether t_c is obtained through τ or from independently available distance and velocity information through division (e.g., Cavallo & Laurent, 1988; Lee et al., 1983; McLeod & Ross, 1983). Very little analysis of the limitations of these two

strategies has been published. In this article I aim to clarify the two strategies and to analyze their success as accounts of human t_c judgements.

First, I examine the tau account. Two questions may be posed with regard to the role of the variable τ in interceptive timing are considered: What are the conditions under which τ is able to play a useful role in timing? In the conditions under which τ may be usefully employed, is it alone sufficient to account for human timing skills, or is additional information required? These questions are addressed in turn in the following sections.

Conditions Under Which Tau Depends on Approach

Tau has been defined by Lee in three ways, which ought to be distinguished to avoid confusion. One definition refers to the focus of expansion of a locomotor optic-velocity field (Lee, 1980), another refers to the relative rate of separation of two points on the image of a moving object (Lee, 1976), and the other refers to the rate of dilation of the image of a moving object or surface patch (Lee & Young, 1985). Consider the first of these: An animal is moving through a rigid environment such that a flow field is imaged on its retina. A convenient representation of this flow is the image-velocity field—the instantaneous projection of the real-world velocities of all visible points in the environment onto an imaging surface. If the organism's eye is not rotating in relation to the environmental frame of reference, this image velocity field will correspond to a projection of what Gibson called the *optic-flow field* and will contain a focus of expansion (Gibson et al., 1955; Longuet-Higgins & Prazdny, 1980). With the retina treated as a plane (but see Appendix A), if a small feature (e.g., a texture element) on a surface being approached by a moving animal is imaged a retinal distance $r(t)$ from the focus of expansion at an instant of time t and moves away from it with a retinal velocity $\dot{r}(t)$ (a dot over a variable represents its temporal derivative), then Lee (1980) defines $\tau(t)$ to be $r(t)/\dot{r}(t)$. Under these conditions, if the animal's velocity remains constant, $\tau(t)$ is equal to the t_c of the texture element with a

I acknowledge the helpful critical remarks made on a previous version of this article by J. T. Todd, W. Schiff, and M. Kaiser. I also thank Dave Lee for comments on the work presented in this article and for much helpful discussion concerning time to contact. Some of the work reported here was conducted when the author was in receipt of a postgraduate studentship from the Medical Research Council of Great Britain.

plane through the eye (strictly through the point of projection) normal to the direction of travel. Because this definition refers to a feature of the velocity field as a whole, it may be called *global tau*, $\tau_g(t)$. Notice that extraction of $\tau_g(t)$ requires detection of the coordinates of the focus of expansion in a retinal-coordinate system.

Now consider the second definition. (Figure 1) shows a surface patch moving in relation to an observation point, O. If the velocity is constant and directly toward O and the angular velocity is zero, then the reciprocal of the rate of dilation of the angle $\alpha(t)$ subtended at O by two points (e.g., the pair of points a, b or the pair c, d) on the surface patch at an instant of time t is equal to the t_c of the surface patch with O (Lee, 1976). The instantaneous rate of dilation of $\alpha(t)$ is $\dot{\alpha}(t)/\alpha(t)$ and can be conceived of as the rate of separation of the images of the two points on a spherical projection surface centered on O. Lee (1976) called the reciprocal of this rate of dilation tau.

Finally, consider the third definition. Under the same assumptions made for the second definition, the reciprocal of the relative rate of dilation of the image of the surface on a spherical projection surface is equal to half the t_c of the surface patch with O. This quantity is equivalent to $\Omega(t)/\dot{\Omega}(t)$, where $\Omega(t)$ is the instantaneous solid angle subtended at O by the surface patch. Lee and Young (1985) defined $\tau(t)$ as the quantity $2\Omega(t)/\dot{\Omega}(t)$, which is equal to t_c under the conditions stated. Note that tau, like time to contact, is an instantaneously defined quantity. This dependency on time has been represented by writing $\tau(t)$; however, throughout the remainder of this article dependency on t is suppressed for brevity.

The latter two definitions are defined locally to the image of an approaching object and may be distinguished from the first definition by calling the quantity so defined *local tau*, τ_L .

In what follows it is necessary to distinguish the two different definitions of local tau. Thus the first (in terms of two points) is denoted by $\tau_L^{(1)}$ and the second (in terms of the image as a whole) is denoted by $\tau_L^{(2)}$ when a distinction is required. (Note that τ_L may also be defined by using a planar projection surface, which leads to slightly different consequences. The definitions that use visual angles, equivalent to use of a spherical projection surface, are adopted here for reasons that are detailed in Appendix A.) This article is primarily concerned with τ_L , which is the quantity that is relevant to interceptive acts like catching and hitting (considered in the following) in which the performer interacts with an object moving in relation to the environment, thus violating the global-rigidity constraint required in the definition of τ_g . In natural scenes, whenever global tau is defined, so is local tau. The converse is clearly not true, however; if an observer is stationary, no global velocity field is defined, so neither is τ_g .

For τ_L to give an actual time to contact, the conditions of constant velocity and direct collision course with the observation point must be met. It is possible, however, that even if these conditions were violated, τ_L may still play a useful role in interceptive timing (this is discussed in the next section). Other conditions must also be met, however, because the image of an object can change size independently of movement toward or away from the observation point. The image of an inflating balloon expands independently of such motion, for example. Therefore, the first constraint that is required if τ_L is to give t_c under conditions of constant-velocity collision is that the approaching object be rigid.

The size of a rigid object's image can change independently of approach, as has been pointed out in the context of timing by Koenderink (1985; see also Koenderink & van Doorn, 1975) and Regan (Regan, 1986; Regan & Beverley, 1979). In the case of direct approach, this may happen when a nonspherical object rotates as it approaches (what happens in oblique approaches is considered in the next section). Consider the situation illustrated in Figure 1. When the surface patch approaches along a collision path and rotates with the illustrated (constant) angular velocity ω , the expression $\Omega = A \cos \theta / R^2$ for the solid angle subtended by the patch may be differentiated with respect to time to yield an expression for the rate of change of solid angle, $\dot{\Omega}$. Dividing the expression for $\dot{\Omega}$ by that for Ω gives an expression for the rate of dilation of the solid angle:

$$\frac{\dot{\Omega}}{\Omega} = \frac{2V}{R} - \dot{\theta} \tan \theta. \tag{1}$$

Thus the rate of dilation is the sum of two terms: The first depends only on the approach ($2V/R$) and is twice the t_c with observation point. The second term depends on the rotation because $-\dot{\theta} = \omega$. Thus $\tau_L^{(2)}$ for a nonspherical object specifies t_c with O only in the absence of rotation. This compromises the utility of $\tau_L^{(2)}$ somewhat. As has been pointed out elsewhere (Regan, 1986; Tresilian, 1990), however, one can get around this problem to some extent. This involves the use of $\tau_L^{(1)}$ rather than $\tau_L^{(2)}$. To see how, consider again Figure 1. The quantity $\tau_L^{(1)}$ defined according to the second of Lee's definitions may be measured for the pair of points a and b or for

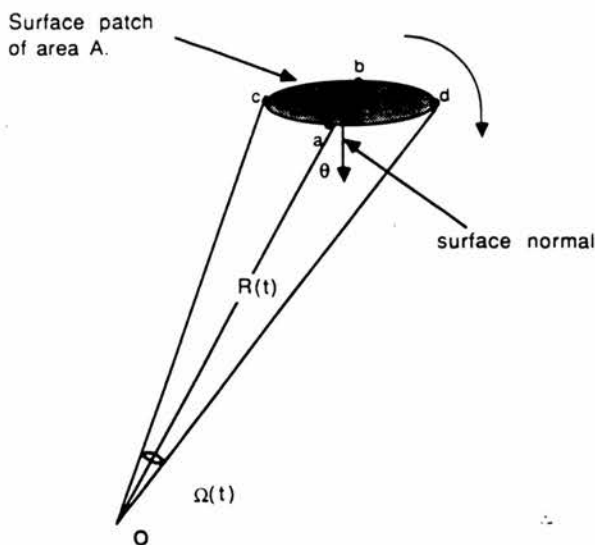


Figure 1. A small surface patch of area A is instantaneously a distance $R(t)$ from a point of observation O and subtends a solid angle $\Omega(t)$ at O. The angle between the surface normal and the line of sight is $\theta(t)$. The patch translates with some (constant) velocity and rotates around an axis through the points a and b with constant angular velocity ω . In the text, dependency on t is suppressed.

the pair of points c and d . When the patch is rotating with velocity ω , $\tau_L^{(1)}$ for the point pair (c, d) has a contribution that depends on the rotation. On the other hand, $\tau_L^{(1)}$ for the point pair (a, b) is independent of the rotation and depends on the approach only: It gives the t_c . Thus, if τ_L (as $\tau_L^{(1)}$) is measured for the dimension perpendicular to that which depends on the object rotation, it specifies t_c . In effect, one is trying to find a method for measuring the image dilation that is due to motion toward (or away from) the observer independent of changes that are due to object rotation. Rotation induces changes in the shape of the image as well as changes in area (size), and for the case described previously it is possible to design a mechanism that automatically measures the appropriate $\tau_L^{(1)}$ and filters out changes due to rotation (Beverly & Regan, 1980).

It is possible for the surface patch illustrated in Figure 1 to have a component of rotation perpendicular to ω (i.e., around an axis along the line joining c and d) such that $\tau_L^{(1)}$ measured for the point pair (a, b) no longer depends on the approach only. Clearly, if ω is zero under these circumstances, then $\tau_L^{(1)}$ for the point pair (c, d) gives t_c . On the other hand, if the surface patch is rotating around both axes simultaneously, then τ_L , as $\tau_L^{(1)}$ or $\tau_L^{(2)}$, fails to give t_c . Notice that for certain nonspherical objects such as American footballs, the situation never gets quite as bad as this in the case of collision approaches—for an American football only rotation along one axis can affect image size. Noncollision trajectories can be more complicated because the angle θ in Figure 1 can change here either because of object rotation (as discussed before) or because of motion of the surface patch in relation to O . I return to this point in the following (see also Appendix A).

Is τ_L Alone Sufficient for Skilled Timing?

Two questions may be considered under this heading: Because τ_L can only specify t_c when the velocity of the approaching object is constant, is it a useful quantity for informing about t_c when the object is accelerating? Because τ_L is only able to specify t_c with the observation point, is it a useful quantity in timing interactions with objects at points some distance from the observation point?

Constant-Velocity Assumption

Lee (1980) introduced the idea that even in the presence of relative acceleration between an animal and an object (or surface), the animal might time an interaction with such an object on the assumption that the velocity is constant. Lee dubbed this the *tau-strategy* (Lee & Reddish, 1981; Lee et al., 1983) because this is the strategy an animal would follow if it were using τ to time interactions with moving objects. To avoid confusion, it should be stressed that Lee's tau-strategy is independent of whether τ is actually used in interceptive timing; for this reason, it might better be called the *constant-velocity strategy* (this name is used throughout this article).

A small but significant body of empirical results support the notion that a constant-velocity strategy is adopted in interceptive timing. First, Lee showed that gannets, when

plunge diving to catch fish, follow such a strategy in timing the retraction of their wings before entering the water (Lee & Reddish, 1981). In another study he presented evidence that human subjects leaping to punch a ball that is falling toward them (thus accelerating under gravity) follow the constant-velocity strategy (Lee et al., 1983). Second, Todd (1981) measured the threshold for detecting differences in t_c in computer-generated images and found that observers were unable to take accelerations into account effectively. Third, several studies indicate that human observers are insensitive to object acceleration perpendicular to the line of sight (Gottsdanker, 1956; Gottsdanker, Frick, & Lockerd, 1961; Runeson, 1975). For example, Gottsdanker et al. (1961) found that smoothly accelerated motion was not well discriminated from unaccelerated motion. This effect was most marked for brief viewing times (less than 1 s)—with longer exposure discrimination improved. Gottsdanker et al. interpreted these results to mean that acceleration is not identified by "direct sensing" (p. 31), which may be taken to mean that information which specifies target acceleration is apparently not being used by the subjects in their judgments. Subject's judgments were consistent with a strategy of discriminating accelerative motion on the basis of the target having different velocities at different times. (Noticing that a target has a different velocity at one time rather than another is not the same as perceiving acceleration.) These latter results provide indirect corroboration of a constant-velocity strategy in interceptive timing by showing that where information about object acceleration might be used, it apparently is not.

In addition, although it can be shown that there is optical information available to specify t_c in linear constant accelerative approaches under a variety of conditions (details of two relevant cases are presented in Appendix B; see also Todd, 1981), this information may be considerably more difficult to extract than information about t_c under the constant-velocity assumption. The expressions that define the information (e.g., see Appendix B) involve a number of different perceptual observables combined in relatively complex ways. Obtaining this information is likely to be more difficult and more prone to error in measurement and computation than obtaining the corresponding information with the assumption of constant velocity. Although the fact that the information is not or cannot be used is no argument in itself, Lee pointed out that the constant-velocity strategy can be used effectively to time interceptions during accelerative approach (Lee et al., 1983) simply and robustly, making it unnecessary to extract the more complicated information, which takes acceleration into account. Although the cited evidence is consistent with a constant-velocity strategy, it is important to note that direct evidence for such a strategy in interceptive timing has only been provided for the case of collision (or approximate collision) with the observer's eyes (Lee & Reddish, 1981; Lee et al., 1983). No direct evidence for the constant-velocity strategy for noncollision trajectories has been published.

Is Timing Based on Tau Alone?

Two cases are discussed here: one in which an object is on a collision course with an observer's eye and the observer

interacts with the object a small distance in front of the eye, and one in which an object is not on a direct collision course but rather passes by a short distance away from the observer, who intends to interact with it while it is within reach. Consider the first case: The optic variable tau specifies t_c of an approaching object with the observation point (the eye). However, if the object is a baseball heading directly toward the head of an outfielder, for example, the fielder will catch it a short distance in front of the eye. Consider the situation represented schematically in Figure 2.

In Figure 2, the time to contact with the eye (Z/V , where V is the speed) at an instant of time is given by the value of τ_L at that time. Time to contact with the interception point, T_i , is given by

$$T_i = \frac{(Z - d)}{V}, \quad (2)$$

which can be written as

$$T_i = \tau_L - \frac{d}{V} \quad (3)$$

Thus, taking τ_L as an estimate of T_i introduces an error equal to d/V . Suppose that the catcher's arm is outstretched such that the distance, d , of the hand from the eye is 50 cm. If the ball were traveling at 8 m/s then the margin of error for a successful catch is about that estimated by Alderson, H. Sully, and D. Sully (1974), 15 ms or so. Under these conditions, the error induced by taking τ_L as an estimate of time to contact is 62.5 ms. Skilled catching with the arm outstretched would therefore be impossible; it would be fumbling and incompetent at best. Skilled timing of the catch would only be possible at distances of less than about 12 cm in front of the eye. This does not accord well with everyday experience of catching skills, and the reader is encouraged to try catching a soft ball falling toward his or her head with outstretched arm. One needs not only τ but information about distance, d , to the interception point and the speed V of the approaching object if timing is to be precise.

The situation just described has some bearing on the theory of braking control proposed by Lee (1976). Lee suggested that the rate of change of τ , that is, its temporal derivative $\dot{\tau}$, can be used to inform about whether a car driver's deceleration is adequate to stop in time to avoid collision with an upcoming obstacle. Lee argued as follows. Through the application of Newton's equations, he deduced that the driver's deceleration (D) is adequate if and only if the distance it will take the car to stop with that deceleration is less than or equal to the current distance (Z) from the upcoming obstacle, that is, if and only if

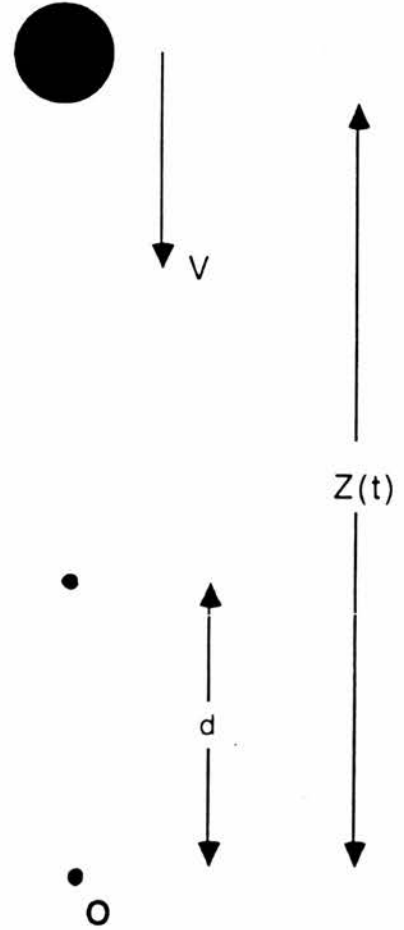
$$\frac{V^2}{2D} \leq Z. \quad (4)$$

From this, Lee deduced that deceleration is adequate if and only if

$$\dot{\tau} \geq -0.5, \quad (5)$$

and he concluded that "a safe braking strategy would therefore consist of the driver adjusting his braking so that $d\tau(t)/dt$

Approaching
object



interception
point, i

Figure 2. A moving object (ball) is approaching a point of observation O on a collision course with speed V (considered constant). It is instantaneously a distance $Z(t)$ from O. The observer requires the time to contact with the interception point, which is a distance d in front of O.

remained a safe value" (Lee, 1980; p. 294). This conclusion that the strategy is safe is unwarranted, however, because no account is taken of the extent of the car body in front of the driver. Suppose that the car extends a distance d in front of the driver; taking this into account requires that Z in Equation (4) be replaced with $Z - d$, the distance of the obstacle from the front of the car. Following the same arguments as Lee yields, instead of Equation (5), the expression

$$\dot{\tau} - \left(\frac{\dot{\tau} + 1}{\tau} \right) \frac{d}{V} \geq -0.5. \quad (6)$$

As Lee has pointed out (personal communication, January 1990), the quantity d/V in Equation (6) is the time to contact with the plane of the driver's eye of the "piece" of road instantaneously next to the front of the car. This t_c is specified by the value of global tau for a texture element lying on this piece of road. Thus, denoting the tau value for the obstacle by τ_o and the tau value for the texture element on the road by τ_n , Equation (6) may be written entirely in terms of optical

variables, as follows:

$$\dot{\tau}_o - \left(\frac{\dot{\tau}_o + 1}{\tau_o} \right) \tau_r \geq -0.5. \quad (7)$$

Thus, in considering the extent of the car in front of the driver, it is not enough simply to use the quantity $\dot{\tau}$ as Lee originally suggested. Adopting Lee's braking strategy entails making errors in braking that are likely to lead to crashes because the driver's tendency is to brake too late. It cannot be concluded, therefore, that Lee's strategy is a safe braking strategy.

What happens if one attempts to use tau as an approximation of t_c when the object is not on a collision course with the eye? The situation represented in Figure 3 is considered. In this case, the relation between τ_L and the illustrated environmental quantities, derived by von Hofsten and Lee (1985), is

$$\tau_L = T_N + S^2/ZV, \quad (8)$$

where T_N is the time to contact with the point of nearest approach. Assume that the observer wishes to intercept the object at the point nearest approach (N). This seems a sensible point to consider because if the object is intercepted in front of N (nearer the object in Figure 3), the error involved in taking τ_L as an approximation to t_c will be slightly less than that at N , whereas if the object is to be intercepted behind N , the error will be slightly greater. The error introduced by using

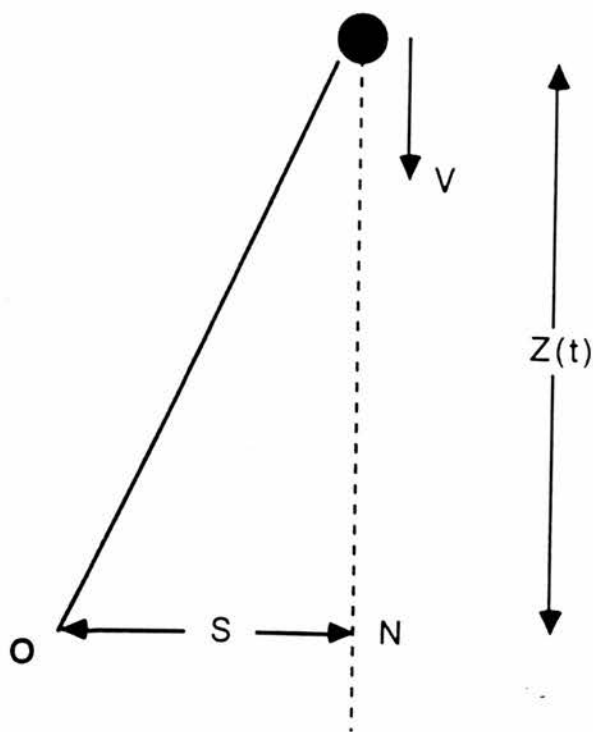


Figure 3. A ball moving in the plane of the page is approaching an observation point O on a miss path (dotted line) with constant speed V . The ball is instantaneously a distance $Z(t)$ from the point, N , on its projected path nearest the point of observation (the point of nearest approach). The t_c of the ball with N is equal to $Z(t)/V$.

τ_L as an approximation to time to contact in this situation is of magnitude S^2/ZV (this is assuming τ_L under spherical projection; see Appendix A). Is this error acceptable? If S is small or V is large, it may be satisfactory. Consider a slip fielder in the game of cricket. Suppose that he or she is faced with the problem of catching a ball traveling at 20 m/s, which will pass to his or her right at a catchable distance, say 50 cm. Suppose further, on the basis of the findings of Whiting and his colleagues (e.g., Sharp & Whiting, 1974, 1975; Tyldesley & Whiting, 1975), that information obtained 300 ms before contact is adequate for reasonably effective catching. The error in taking τ_L as an estimate of T_N may be calculated to be about 2 ms. Thus the tau approximation supplies very precise information indeed under these conditions. It is unlikely that a biological vision system could calculate τ_L accurately enough to exploit this precision, but it may not need to—a temporal accuracy of 5–10 ms is probably required to catch effectively in the cricket slips (Lee & Young, 1985).

When the ball is moving at the more moderate speed of 6 m/s, the error increases to 23 ms ($S = 50$ cm), but this may still be just about good enough to achieve reasonable catching success (Alderson et al., 1974; Sharp & Whiting, 1974). It is at slower speeds that taking tau as an approximation begins to look really suspect. When the object to be caught moves at a speed of 4 m/s, the error ($S = 50$ cm) increases to 52 ms, and at 1 m/s it is 800 ms. Indeed, with ball speeds of less than 1 m/s von Hofsten (1983) found that the catching skills of very young infants were far more accurate than could be accounted for by a tau approximation strategy.

It may be concluded that at very fast object speeds a tau approximation strategy is effective, but it is less so for slow and moderate ball speeds. At these speeds some other, more precise, information is required. It has recently been shown that by assuming the constant-velocity approximation, the required information is perceptually available (Tresilian, 1990). In Appendix B the derivation of this information is given. (I further show that equivalent information is available without the assumption of constant velocity, though it is unlikely that human observers use this information as discussed previously). The argument in Appendix B shows that the time to arrival (t_p) at any point p on the upcoming path of an approaching object is given by

$$t_p = \frac{\tau_L}{1 + \tau_L u \cot \psi}, \quad (9)$$

where u is the instantaneous rate of change of direction of the moving object measured at the eye and ψ is the angle between the instantaneous position of the approaching object and the point on its upcoming path (see Appendix B for more details).

Note that because the information described by Equation (9) involves the τ_L variable, it is subject to the limitations described earlier. See Figure 1: If the object is not spherical, is not rotating ($\omega = 0$), and passes the observer on a miss path, then the image of the surface patch will change size not because the distance from O changes but because the angle

¹ Because $Z/V = 0.3$ s and $V = 20$ m/s, $Z = 0.3 \times 20$. The error S^2/ZV is equal to $0.5^2/0.3 \times 20 \times 20$, which comes out to 2.08 ms.

between the surface normal and the line of sight changes. Thus τ_L defined as the inverse of the rate of dilation of the object's image (i.e., $\tau_L^{(2)}$) will not give the quantity that is required in Equation (9) (see Appendix B). The required quantity is given by $\tau_L^{(1)}$ measured along the direction perpendicular to the direction of object motion. If the object is rotating around an axis through (a, b) while approaching along a miss path, however, $\tau_L^{(1)}$ cannot specify the quantity required in the definition of Equation (9) whatever direction it is measured along. The following considerations suggest another way to measure τ_L so that it does specify the required quantity.

Suppose that an irregularly shaped object is rotating but is not changing its distance from the observer. The image will continuously change in size (and shape), but the average image size over complete revolutions of the object will be constant. If this object is now considered to approach the observer along a miss path while rotating, and the average area (or average solid angle subtended at the observation point by the object) of the image over the i th complete revolution is \bar{A}_i , then the rate of change of the area can be approximated by

$$V^* = \frac{\bar{A}_{(i+1)} - \bar{A}_i}{\Delta t} \quad (10)$$

where Δt is the time taken to make a measurements of average area (i.e., the period of the object's revolution). The quantity τ_L might then be roughly approximated by $2V^*/\bar{A}_i$. It is by no means easy to come by, however. For example, how does one ensure that the area has been averaged over a complete revolution? Information about how the object is rotating is required if appropriate estimates are to be made. Fortunately, this situation is not so complicated for a binocular observer, and it may be that monocular human observers find it very difficult to deal with the case that has just been considered (though there seems to be no direct empirica evidence on this matter). As shown elsewhere (Tresilian, 1990), a binocular observer has access to t_c information that does not require the quantity τ_L and consequently does not suffer from the problems involved in using this quantity. This binocular t_c information has nothing to do with computing distance and velocity but rather has the form of Equation (9). This information cannot be used when the object is approaching the observation point directly, but this is not a severe limitation. A monocular special case of this information that illustrates these features is described in the next section.

Distance Divided by Velocity Account of t_c Perception

It is possible that an observer could obtain t_c information not from a relation that specifies it directly without need for further computation (such as that discussed in the previous section) but from perceived distance and velocity information. If the observer can obtain information about the distance of the moving target from the interception point and its velocity, then (assuming that velocity remains constant) t_c of the target with the interception point could be obtained by dividing this distance by this velocity. Discussion in this section continues

to assume constant velocity for simplicity (the argument I develop remains essentially unchanged if acceleration is taken into account: this ought to be clear from Appendix B). Although this method for obtaining t_c information is frequently cited as an alternative to a tau-based account, no fully worked out distance divided by velocity scheme is to be found in the literature. What needs to be specified is precisely what perceptual information about real-world distances and velocities might be used to compute t_c .

There is evidence that t_c can be estimated when no absolute distance and velocity information is available (Schiff & Dewteler, 1979; Todd, 1981), which may be taken as evidence for the use of a tau-based strategy. This does not establish that when such information is available it is not used to compute t_c ; neither does it exclude the possibility recently suggested by Schiff and Oldak (1990) that in some circumstances tau-based strategy is used and in others distance divided by velocity strategy is used. With regard to this latter suggestion, there are certain situations in which tau information is unlikely to be of much use. For example, consider the early games of computer "tennis" in which a small square "ball" composed of a few pixels moves rapidly across the video screen, and the player intercepts the motion of this moving square with a simulated "bat" several pixels across. The value of τ_L for this moving square on the retina of a player a meter from the screen is vanishingly small. This means that the information described by Equation (9) might be unusable in timing hits of the "ball" with the "bat." In these circumstances a distance divided by velocity strategy might seem to be required. It is easy to show, however, that t_c information may be obtained simply without the need to obtain distance and velocity information first. The geometry of the situation is represented schematically in Figure 4 from which the relation $t_c = w/\dot{w}$ is easily obtained.

The quantity w/\dot{w} is a retinal distance divided by its rate of change and is thus identical in form to tau—it is the reciprocal of a rate of dilation. Thus, in this special case a monocular observer can obtain t_c information without first having to obtain information about the distance W and the speed V and without having to detect any changes in image size. To avoid any confusion here, note that the use of the quantity w/\dot{w} to obtain t_c is not a distance divided by velocity strategy because both w and \dot{w} are observables of the projection, not environmental quantities.

Because tau has been defined already in several different ways, it might be appropriate to call w/\dot{w} tau as well. This may be done consistently by defining tau in general to be the reciprocal of the rate of dilation of a perceptual quantity; this quantity need not be visual, it may be haptic (e.g., see Carello, Kugler, & Turvey, 1985; Kelso, 1986) or acoustic (Lee, in press). Any quantity such as τ_R , τ_L (as $\tau_L^{(1)}$ or $\tau_L^{(2)}$), or w/\dot{w} described earlier is then an example of a tau-type variable. Distinctions between different taus can readily be made by using subscripts, as has been done in this article.

Of course, the possibility remains that distance and velocity are used to obtain t_c information when they are available. A variety of possibilities exist as to the source of such information. A monocular observer may use optical information about distance and velocity (described in the following) or

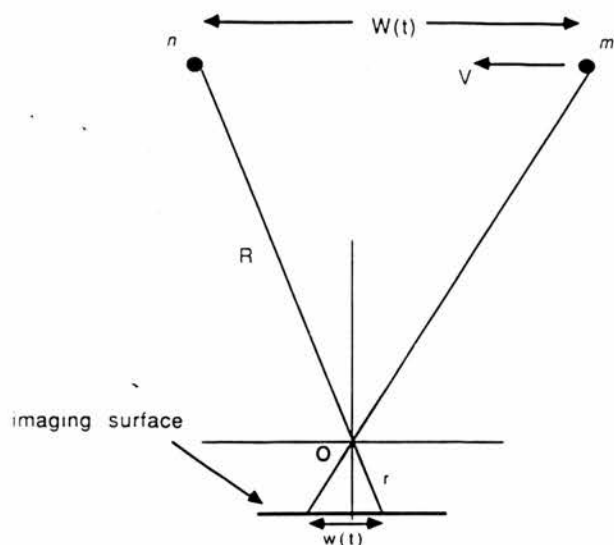


Figure 4. Optic geometry of time to contact for two points lying parallel to the image plane. (The point m is moving and n is fixed. The velocity of the moving point is constant and along the line joining the two points. By similar triangles $W(t)/R = w(t)/r$, which when differentiated with respect to time gives $V/R = \dot{w}(t)/r$. Division of the first of these relations by the second gives the equation given in the text, where t_c is the time to collision of the two points m and n and is equal to $W(t)/V$. As usual, dependency on t is suppressed in the text. Note that the use of a planar projection surface in this case ensures that the expression for time to contact has a particularly simple form: Adoption of a spherical model complicates matters. For further discussion of this general point, see Appendix A.)

perhaps ocular accommodation cues. A binocular observer may also use binocular disparity information (appropriately scaled) or binocular convergence angle (again appropriately scaled) to obtain information about the objective distance of the target. Several researchers have interpreted experimental results as indicating the use of a distance-velocity strategy in certain conditions (e.g., Cavallo & Laurent, 1988; Groeger & Brown, 1986; McLeod & Ross, 1983). The data, however, do not support such an interpretation, a conclusion that is discussed fully in the next section. In a previous article (Tresilian, 1990), I identified an implicit assumption in a monocular formulation of the distance divided by velocity account. I briefly reiterated this argument in the following and extend it to cover any account for computing t_c on the basis of the objective distance and velocity of a moving target.

In Appendix A expressions are derived (after Todd, 1981) for the (horizontal) distance Z of a moving target from a monocular observer and its horizontal velocity V_z , both scaled in terms of the object's size, B . For reference, the scaled distance is given by $Z/B = 1/b$ and the scaled velocity by $V_z/B = \dot{b}/b^2$, where b is the length of the object's image on a planar projection surface. Notice that because b/b is the reciprocal of the rate of dilation of the target's image, it may be denoted by τ_L . The expression for target velocity can then be written as $V_z/B = 1/b\tau_L$, which indicates that τ_L may play a role in the extraction of quantities other than t_c . It has thus been shown that if a monocular observer has prior knowl-

edge about the objective size B , then he or she can obtain information about the object's distance away and its velocity. Dividing the expression for distance by that for velocity, one obtains the following expression for Z/V_z , a time to contact:

$$Z/V_z = b/\dot{b} = \tau_L. \quad (11)$$

This derivation of an expression for t_c can be read in two ways: either as a purely formal argument showing that t_c information is visually available or as an account of how the visual system might actually compute t_c . Reading in the latter sense gives the following account for t_c perception by a monocular observer: The observer's visual system computes distance and velocity according to the two expressions given before (it "knows" about the object's size, B) and then divides these two perceived quantities to obtain t_c . The former account says that the observer's visual system simply measures τ_L . It does not need to obtain it by way of the roundabout method of measuring $1/b$ and $1/b\tau_L$, scaling these two quantities with respect to prior information about the object's size, and then dividing the results of these two processes to get the same result as would be obtained by measuring τ_L .

Such a distance divided by velocity account clearly begs the following question: Why should the visual system ever take the roundabout distance divided by velocity route (which is probably more prone to error) to t_c information when the same information may be extracted in a much simpler way? It is difficult to see what the rationale for such a route would be. This problem is not only a feature of the monocular formulation of the distance divided by velocity strategy presented earlier; a similar problem arises for any such strategy, because to obtain information about the distances and velocities required to compute t_c , the observer must scale the relative depth and velocity information present in the stimulus. For example, in the case of an object some distance away but directly in front of the observer, binocular disparities and convergence angles must be scaled by the interocular separation if the distance away is to be obtained from either of these sources. Thus any distance divided by velocity strategy first scales stimulus information and then undoes this scaling to get t_c information. Although a perceiver might be driven to use a distance divided by velocity strategy under unusual and impoverished conditions, without some guiding rationale it is difficult to argue that such a strategy would be used in normal circumstances, in which it is an uneconomic method for computing t_c . Nonetheless, as noted before, certain empirical results have been interpreted as evidence for the use of a distance divided by velocity strategy by human observers. In the next section I consider the interpretation of empirical results.

Empirical Studies

It is frequently argued that if it could be shown empirically that human observers used more than tau alone to judge t_c , this would go against a tau-based account of the perception of t_c and tend to support instead some kind of computational account that is presumably based on distance and velocity (e.g., Cavallo & Laurent, 1988; Groeger & Brown, 1986;

McLeod & Ross, 1983; McLeod et al., 1985). The earlier discussion demonstrates that evidence for factors other than tau in t_c judgments is precisely what is expected from a tau-based account extended to deal with the real-world requirements of interceptive timing. In this section, studies by Cavallo and Laurent (1988) and by Schiff and Detwiler (1979) are discussed, both of which have been interpreted as bearing on whether a distance divided by velocity or tau-based strategy is used to obtain t_c information. Note that both of these studies used a t_c judgment task that required subjects to press a button at the moment a contact was judged to have occurred. Such a task is clearly neither a precise interceptive act like a catch nor an avoidative action like braking. It is thus unclear whether the results reported in the studies reviewed here are relevant to natural timing behaviors. Nevertheless, this does not preclude use of the results to argue for or against a tau or distance divided by velocity strategy in t_c estimation. In what follows, however, I argue that the results cannot be interpreted as distinguishing the two strategies.

Cavallo and Laurent (1988) reported data that was collected from subjects who were driven at a target in an automobile. Shortly before contact with the target, the subjects' vision was occluded, and they were required to press a button at the moment they expected the car to collide with the target. A major finding was that the extent of the subjects' visual field had a significant effect on t_c estimation for inexperienced but not for experienced drivers. The larger the field of view, the better the estimates made by the inexperienced drivers. Cavallo and Laurent proposed that if tau information alone were being used, then one would expect the results to be independent of the size of the visual field. This is perhaps true for local tau but false for global tau because as shown theoretically by Koenderink and van Doorn (1987), the accuracy with which the focus of expansion can be localized depends on the field of view: In general, the larger the field of view, the better it will be localized.

Cavallo and Laurent observed, however, that even if a case could be made for an enhancing effect of a wide visual field on the pickup of tau, this is expected to be independent of driving experience. They concluded that the visual-field effect occurs because beginners assessed speed as a separate parameter, and it may thus be interpreted as using a distance divided by velocity strategy, whereas the performance of the experienced drivers is consistent with a tau-based strategy. This explanation of the visual-field effect raises the question of why the novice drivers used a completely different source of information from the experienced drivers. The analysis of the tau account presented earlier suggests two alternative explanations that do not suffer from this difficulty. First, it is possible that driving experience makes the extraction of global tau more robust against decreases in the size of the visual field. Second, returning to the analysis presented earlier with reference to Figure 2, it is clear that in order to judge accurately when the front of the car will hit an obstacle, the velocity of the car and the extent of the car in front of the driver need to be considered; see Equation (3). It is then possible to argue that when the field of view is restricted and vision of the hood is occluded, the driver no longer has access to information about the length of the hood. It would not be

surprising if the experienced drivers had learned about hood lengths through extensive experience and could take the hood into account without actually having to see it. In contrast, novice drivers, are less likely to be able to consider the length of the hood effectively without seeing it. It might therefore be expected that the difference between novice and expert drivers would decrease as the speed of approach to the target increased because the error induced by failing to consider the hood length is negligible at high speeds. Thus the differential effects of field of view would only be evident at rather slow speeds, given the precision of t_c estimation shown by subjects in this kind of task (note that this is not a prediction of the first explanation).

Schiff and Detwiler (1979) presented subjects with short animated film sequences showing a black square directly approaching the camera position. The background of the approaching square was divided into two regions (terrain and sky) by a thin black "horizon" line. Either both terrain or sky were plain white, or one was white and the other was covered in a grid of squares. Schiff and Detwiler claimed that the grid terrain provided subjects with enhanced distance and "distance change" (velocity) information. There was no absolute distance or velocity information in Schiff and Detwiler's stimuli, however, and such information is required for the distance divided by velocity strategy. The reason for the absence of absolute distance and velocity information is clear: there was nothing in terms of which the relative distance and velocity present in the stimuli could be scaled. The subjects were watching a two-dimensional film in which there was apparently no means by which they could scale the scene being viewed—it could be a small scene nearby or a large one a long way off. Thus, the results of this study can shed no light on whether tau or distance divided by velocity is used to obtain t_c information. The data do indicate that subjects are able to use image expansion (in the form of τ) when both distance and velocity information are unavailable.

Conclusions

In this article, several limitations of both tau-based and distance divided by velocity accounts of the perceptual source of t_c information have been discussed. These limitations are usually either overlooked or left implicit in discussions of perceptual timing. In particular, it was shown that an account of timing on the basis of tau alone has significant problems in accounting for the skill and flexibility shown by normal human subjects. More general information has been described that can overcome these problems (Tresilian, 1990). This involves variables other than tau, which indicates that empirical demonstration that factors other than tau are important in t_c judgment cannot be taken simply as support for the alternative distance divided by velocity account. This alternative assumes that the perceptual systems responsible for the extraction of t_c information first scale relative depth and velocity information to obtain objective distances and velocities and then effectively undo this scaling operation to obtain t_c . Any distance divided by velocity account needs to provide a justification for this step or if possible show that it can somehow be avoided while retaining the essence of the strat-

egy. In addition, before empirical study of which strategy is used under what circumstances can be usefully conducted, a detailed formulation of the distance divided by velocity account is required, which leads to predictions different from those of the tau-based account.

Two different types of optical tau were distinguished: a global variable that required the focus of expansion of the translational component of an image-velocity field for its definition and a local variable that is independent of detection of a focus of expansion. Two versions of local tau can be defined, with each having different consequences for perception of t_c . It was noted (Appendix A) that tau defined for a planar model of projection specifies a different t_c from that specified by tau defined for a spherical model of projection and furthermore that use of tau defined for planar projection requires that a Cartesian coordinate system be established for its measurement.

The local tau variable's limitations were discussed in detail. It was shown that there are situations in which τ_L may be very difficult to extract either because it is very small (the image size of the moving object changes very little, such as in some video games) or because the moving object is nonspherical and moving with rotation on a trajectory that bypasses the observer. In these conditions an observer has access to t_c information that does not involve τ_L or entail following a distance divided by velocity strategy. Information of this kind suitable for the video-game case was described earlier. Such information contains a quantity that is the reciprocal of a rate of image dilation, and according to a suitably general definition is an instance of a tau-type quantity.

Although throughout this article I have referred to the account of t_c perception that extends Lee's original tau theory as a tau-based account, it is not necessarily any more tau-based than the distance divided by velocity strategy. I have treated tau here as an observable of the optic projection, which appears in equations defining information about t_c but also in expressions for information about target velocity (see the *Is τ_L Alone Sufficient for Skilled Timing?* section) and acceleration (see Appendix B). What distinguishes the two accounts is not the use of the variable tau per se but whether t_c information is computed directly from observables of the optic projection or from information about real-world quantities that has already been computed from observables of the projection.

References

- Alderson, G. K. W., Sully, H., & Sully, D. (1974). An operational analysis of a one-handed catching task using high speed photography. *Journal of Motor Behaviour*, 6, 217-226.
- Beverley, K. I., & Regan, D. M. (1980). Visual sensitivity to the shape and size of a moving object: Implications for models for object perception. *Perception*, 9, 151-160.
- Carel, W. L. (1961). *Visual factors in the contact analogue* (Publication No. R61ELC60). Ithaca, NY: General Electric Company Advanced Electronics Center.
- Carello, C., Kugler, P. N., & Turvey, M. T. (1985). The informational support for upright stance. *Behavioural & Brain Sciences*, 8, 151-152.
- Cavallo, V., & Laurent, M. (1988). Visual information and skill level in time-to-collision estimation. *Perception*, 17, 623-632.
- Cavallo, V., Laya, O., & Laurent, M. (1986). The estimation of time-to-collision as a function of visual stimulation. In A. G. Gale, M. Freeman, C. Haselgrave, P. Smith, & S. Taylor (Eds.), *Vision in vehicles* (pp. 365-374). Amsterdam: Elsevier.
- Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology*, 68, 372-385.
- Gottsdanker, R. M. (1956). The ability of human operators to detect the acceleration of target motion. *Psychological Bulletin*, 53, 477-487.
- Gottsdanker, R. M., Frick, J., & Lockerd, R. (1961). Identifying the acceleration of visual targets. *British Journal of Psychology*, 52, 31-42.
- Groeger, J. A., & Brown, I. D. (1986). Motion perception is not direct with indirect viewing systems. In A. G. Gale, M. Freeman, C. Haselgrave, P. Smith, & S. Taylor (Eds.), *Vision in vehicles* (pp. 000-000). Amsterdam: Elsevier.
- Hoyle, F. (1961). *The black cloud*. Harmondsworth, England: Penguin.
- Kelso, J. A. S. (1986). Pattern formation in speech and limb movements involving many degrees of freedom. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. ■-■). New York: Springer-Verlag.
- Koenderink, J. J. (1985). Space, form and optical deformations. In D. Ingle, M. Jeannerod, & D. Lee (Eds.), *Brain mechanisms and spatial vision* (pp. 31-58). Dordrecht, The Netherlands: Martinus Nijhoff.
- Koenderink, J. J., & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22, 773-791.
- Koenderink, J. J., & van Doorn, A. J. (1987). Facts on optic flow. *Biological Cybernetics*, 56, 255-260.
- Laurent, M., Dinh Phung, R., & Ripoll, H. (1989). What visual information is used by riders in jumping? *Human Movement Science*, 8, 481-501.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, 5, 437-459.
- Lee, D. N. (1980). Visuo-motor coordination in space-time. In G. E. Stelmach & J. Requin (eds), *Tutorials in motor behaviour* (pp. 281-285). Amsterdam: North-Holland.
- Lee, D. N. (in press). Getting around with light or sound. In R. Warren & A. H. Wertheim (Eds.), *The perception and control of self motion* (pp. ■-■). Hillsdale, NJ: Erlbaum.
- Lee, D. N., & Reddish, P. E. (1981). Plummeting gannets: A paradigm of ecological optics. *Nature*, 293, 293-294.
- Lee, D. N., & Young, D. S. (1985). Visual timing of interceptive action. In D. Ingle, M. Jeannerod, & D. N. Lee (Eds.), *Brain mechanisms and spatial vision* (pp. 1-30). Dordrecht: Martinus Nijhoff.
- Lee, D. N., Young, D. S., Reddish, P. E., Lough, S., & Clayton, T. (1983). Visual timing in hitting an accelerating ball. *Quarterly Journal of Experimental Psychology*, 35A, 333-346.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London*, B208, 385-397.
- McLeod, P., McLaughlin, C., & Nimmo-Smith, P. (1985). Information encapsulation and automaticity: Evidence from the visual control of finely timed actions. In M. Posner & O. Marin (Eds.), *Attention & performance XI* (pp. 391-406). Hillsdale, NJ: Erlbaum.
- McLeod, R. W., & Ross, H. E. (1983). Optic flow and cognitive factors in time-to-collision estimates. *Perception*, 12, 417-423.
- Michaels, C. F., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, NJ: Prentice-Hall.
- Purdy, W. C. (1958). *The hypothesis of psychophysical correspondence*

- in space perception. Unpublished Ph.D. thesis, Cornell University. (University Microfilms No. 58-5594)
- Regan, D. M. (1986). Visual processing of four kinds of relative motion. *Vision Research*, 26, 127-145.
- Regan, D. M., & Beverley, K. I. (1979). Binocular and monocular stimuli for motion in depth: Changing disparity and changing size feed the same motion in depth stage. *Vision Research*, 19, 1331-1342.
- Runeson, S. (1975). Visual prediction of collision with natural and unnatural motion functions. *Perception & Psychophysics*, 18, 261-266.
- Schiff, W., & Detwiler, M. L. (1979). Information used in judging impending collision. *Perception*, 8, 647-658.
- Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology: Human perception and performance*, 16, 303-316.
- Sharp, R. H., & Whiting, H. T. A. (1974). Exposure and occluded duration effects in a ball-catching skill. *Journal of Motor Behaviour*, 6, 139-147.
- Sharp, R. H., & Whiting, H. T. A. (1975). Information processing and eye movement behaviour in a ball-catching skill. *Journal of Human Movement Studies*, 1, 124-131.
- Soloman, H. J., Carello, C., & Turvey, M. T. (1984). Optical flow: the visual support for skilled actions. In W. Straub & J. Williams (Eds.), *Cognitive sport psychology* (pp. 129-139). Lansing, MI: Sport Science Associates.
- Todd, J. T. (1981). Visual information about moving objects. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 795-810.
- Tresilian, J. R. (1990). Perceptual information for the timing of interceptive action. *Perception*, 19, 223-239.
- Turvey, M. T., & Carello, C. (1986). The ecological approach to perceiving-acting. *Acta Psychologica*, 63, 133-155.
- Turvey, M. T., & Kugler, P. N. (1984). The ecological approach to perceiving and acting. In H. T. A. Whiting (Ed.), *Human motor actions: Bernstein reassessed* (pp. 373-412). Amsterdam: North-Holland, 1984.
- Tyldesley, D. A., & Whiting, H. T. A. (1975). Operational timing. *Journal of Human Movement Studies*, 1, 172-177.
- von Hofsten, C. (1983). Catching skills in infancy. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 75-85.
- von Hofsten, C., & Lee, D. N. (1985). Dialogue on perception and action. In W. H. Warren & R. E. Shaw (Eds.), *Persistence and change* (pp. 231-242). Hillsdale, NJ: Erlbaum.

Appendix A

PLANAR MODEL OF THE IMAGING SURFACE

Assume the situation illustrated schematically in Figure A1: A small surface patch approaches the imaging system without rotation. The following relation is a direct consequence of the geometry (dependency on time has been suppressed for brevity as usual):

$$\frac{Z^2}{1} = \frac{A}{\alpha} \quad (\text{A1})$$

which may be differentiated with respect to time to give

$$0 = 2ZZ' + Z^2\alpha' \quad (\text{A2})$$

By dividing (A2) by (A1), rearranging, and writing V_Z for $-Z'$, one obtains

$$\frac{Z}{V_Z} = \frac{2\alpha}{\alpha'} \quad (\text{A3})$$

According to Lee and Young's (1985) definition, the right side of (A3) is τ_L defined on a planar projection surface. Z/V_Z is the time to contact of the surface with a plane parallel to the image plane passing through O if the velocity remains constant. Note that this t_c is not generally the time to nearest approach to the point of observation. This fact is illustrated in Figure A1: if the object is moving along the path shown by the dotted line, the t_c given by Equation (A3) does not correspond to t_c with the point of nearest approach (marked N).

It is a geometrical theorem that projective relations on a projection surface of one shape are related to equivalent relations on a projection surface of another (topologically equivalent) shape by a one-to-one transformation. It does not follow, however, that the projection model one adopts for the eye is without consequence for the relevance of the results obtained to natural vision. To see this, consider the angular geometry (equivalent to spherical projection) in Figure A1. Consider the case of small Ω ; we may write

$$Z = R \cos \xi.$$

This can be differentiated with respect to time to give

$$-V_Z = R \cos \xi - R\xi \sin \xi. \quad (\text{A5})$$

Dividing (A5) by (A4) yields

$$-V_Z/Z = R/R - \xi \tan \xi. \quad (\text{A6})$$

Now, as the surface patch moves the solid angle changes because the angle between the surface normal and the line of sight (θ) changes. Thus the rate of dilation of Ω (or the rate of dilation of the image of the patch on a spherical projection surface) depends on both approach and the change in θ . The effect of the latter is confined to dilation along the Y direction in Figure A1. The rate of dilation along the X direction depends only on the translational motion of the surface patch. In fact, if the extent surface patch along the X direction subtends an angle α at O, then the reciprocal of the rate of dilation of this angle, α/α' (which is τ_L as defined by Lee, 1976), is equal to $-R/R'$ (for small α). Thus, substituting τ_L for $-R/R'$ in (A6) and rearranging gives

$$\frac{Z}{V_Z} = \frac{\tau_L}{1 + \tau_L \xi \tan \xi}. \quad (\text{A7})$$

Thus the quantity Z/V_Z (the t_c with the y axis in Figure A1), although specified by τ_L that is defined on a planar projection surface, is not specified by τ_L that is defined in angular terms (i.e., on a spherical projection surface).

The two taus specify different real-world quantities because τ_L defined on a planar projection surface involves an explicit Cartesian coordinate system (O X Y Z in Figure A1), whereas τ_L defined in angular terms is independent of any coordinate system (it is related to the divergence of the angular velocity field, which is a coordinate-free quantity; e.g., see Koenderink, 1985). To measure τ_L on a planar projection surface of Figure A1, the O X Y Z coordinate system needs to be established in some way. The quantity ξ relates R to the O X Y Z coordinates. In the text τ_L is defined in angular terms because a coordinate-free definition is the most general and is closest to the

18 pts #

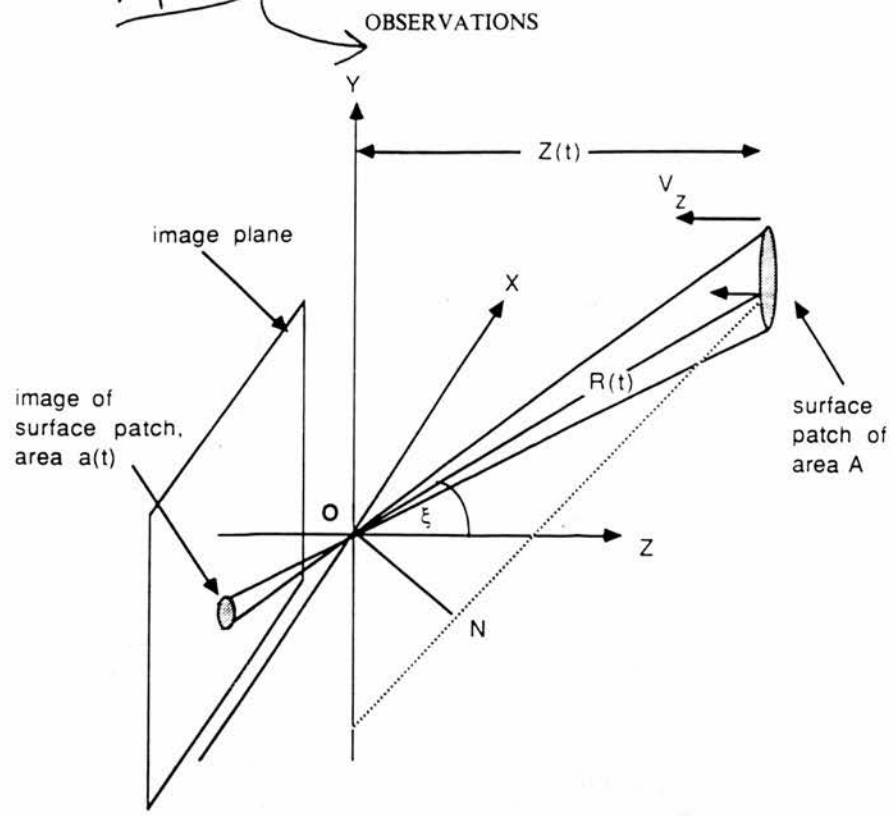


Figure A1. Schematic imaging system with planar projection surface unit distance behind the observation point O. (The Cartesian coordinate system OXYZ is considered fixed in relation to the environmental frame of reference and the small planar surface patch is oriented parallel to the image plane and is instantaneously a distance R(t) from O, which has a component Z[t] in the Z direction. The velocity of the patch lies in the ZY plane and has Z component V_Z [considered constant]. If the path of the object is that shown by the dotted line, the quantity Z[t]/V_Z is the t_c with the Y axis not with the point of nearest approach to O [marked N].)

retinal quantity obtained when the eye follows a moving object with pursuit movements.

Todd (1981) considered a two-dimensional "slice" through the planar projection of Figure A1. Thus the object reduces to a line segment joining two points on the surface patch (say p and q in Figure A1). If the length of this segment is B and the length of its image is b, then one may write

$$Z/B = 1/b \tag{A8}$$

and

$$V_Z/B = b/b^2; \tag{A9}$$

(A8) is obtained from similar triangles, and (A9) is obtained by differentiating (A1) with respect to time. (These two equations appear in the Distance Divided by Velocity Account of t_c Perception section.)

Appendix B

ACCELERATIVE APPROACHES

Consider a spherical object on a collision course with a point of observation O. At an instance of time t, let the distance between the object and O be Z(t), the relative speed by V(t), and the signed magnitude of the relative acceleration A(t) (assumed to be constant). Let the rate of change of direction of the object in relation to O at time t be zero. Dependency on t is suppressed in what follows. Assuming that the expression Z/V = τ_L is valid, it may be differentiated with respect to time to give

$$\dot{\tau}_L = \frac{-V^2 - ZA}{V^2} \tag{B1}$$

Lee et al. (1983) used the following standard formula for distance traveled under constant acceleration to obtain an expression for t_c during a linear accelerative approach:

$$Z = Vt_c + \frac{1}{2} At_c^2 \tag{B2}$$

By treating this as a quadratic in t_c and taking the positive root, the expression derived by Lee et al. for the time to contact of the moving object with O under the conditions described before is

$$t_c = \frac{\tau_L [1 - \sqrt{-2\dot{\tau}_L - 1}]}{(1 + \dot{\tau}_L)} \tag{B3}$$

This shows that for accelerative approach, time to contact is optically

OBSERVATIONS

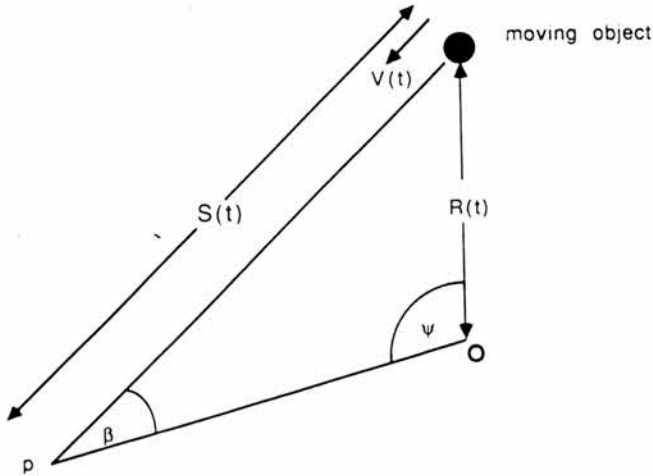


Figure B1. Optic geometry of time to contact. (A spherical object [a ball] is moving in the plane of the diagram [this is a reasonable idealization because in the real world objects tend to have planar motion]. It is moving along a path shown by the dotted line with an instantaneous speed $V(t)$. At the instant of time illustrated, the ball is a distance $S(t)$ from a point p on its upcoming path and a distance $R(t)$ from O .)

specified by an expression considerably more complicated than τ_L alone, which specifies t_c in the same circumstances but with V constant.

Now, consider the geometry illustrated in Figure B1. By assuming constant object velocity and following Tresilian (1990), one can apply the sine rule to obtain (suppressing dependency on t as usual)

$$\frac{R}{\sin \beta} = \frac{S}{\sin \psi} \tag{B4}$$

Differentiating (B3) with respect to time, substituting t_p (the time to contact of the moving object with the point p) for $-S/\dot{S}$ and τ_L for $-R/\dot{R}$ (which is true provided that the conditions discussed earlier hold), and denoting $-\dot{\psi}$ by u (the object's rate of change of direction as seen from O) yields, with rearrangement, Equation (9) given earlier: $t_p = \tau_L / (1 + \tau_L u \cot \psi)$. This expression shows that the t_c with any point p on the upcoming path of the moving object is specified in terms of variables potentially available from the optic projection (full discussion of the conditions under which these quantities are

available is given in Tresilian, 1990). Relaxing the constant-velocity constraint means that Equation (B3) can be differentiated twice and simplified to yield

$$\begin{aligned} \frac{2\dot{V}}{S} - 2\dot{\psi} \cot \psi - \frac{\dot{\psi}^2 S}{V} - \frac{2\dot{R}}{R} + \frac{2S\dot{R}\dot{\psi} \cot \psi}{VR} \\ + \frac{S\dot{\psi} \cot \psi}{V} + \frac{SR\ddot{R}}{VR^2} = \frac{A}{V^2} \end{aligned} \tag{B5}$$

where V is the instantaneous speed of the object. Notice that the quantity $-\dot{R}/R$ is the reciprocal of τ_L (which will be denoted by ρ) and that S/V is t_p , for which an expression in terms of optic variables has already been derived (see Equation [9]). An expression for \dot{R}/R in terms of τ_L and $\dot{\tau}_L$ can be obtained by differentiating $\tau_L = -R/\dot{R}$ with respect to time and substituting for $-R/\dot{R}$ in the resulting expression. $-\dot{R}/R = (\dot{\tau}_L + 1)/\tau_L$. Equation (B5) can be rewritten completely in terms of optic variables as follows (writing t_p for the right side of Equation [9]):

$$\begin{aligned} \frac{A}{V^2} = \frac{2}{t_p} - 2\dot{\psi} \cot \psi - \dot{\psi}^2 t_p + 2\rho - 2\rho t_p \dot{\psi} \cot \psi \\ + t_p \ddot{\psi} \cot \psi - t_p \rho^2 (\dot{\tau}_L + 1). \end{aligned} \tag{B6}$$

To obtain an expression for time to contact under accelerative approach, one finds an expression for t_c from Equation (B2) by applying the standard formula for the roots of a quadratic, noting that in this case t_c is interpreted as the time to contact under accelerative approach with the point p . Thus,

$$t_c = \frac{-1 + \sqrt{1 + 2(A/V)(S/V)}}{(A/V)} \tag{B7}$$

Because expressions for both A/V and S/V have been derived that gives these quantities in terms of optic variables (Equations [B6] and [9], respectively), the time to contact of the moving object with any point p is specified optically under conditions of linear accelerative approach. It is clear that the relevant expression is a complex relation (it is not written out here, but if desired it can be obtained by simple substitution into [B7]) among several optic quantities (two of these, $\ddot{\psi}$ and $\dot{\tau}_L$, are second temporal derivatives) and considerably more complicated than the corresponding expression, which assumes constant velocity (Equation [9]).

Received March 22, 1990

Revision received November 30, 1990

Accepted September 6, 1990. ■

8